

UNIVERSIDADE FEDERAL DO PARANÁ

CAMILA DE OLIVEIRA

IDENTIFICAÇÃO E CORREÇÃO DA BANDA BRILHANTE EM DADOS DE
RADAR METEOROLÓGICO

CURITIBA

2014

CAMILA DE OLIVEIRA

IDENTIFICAÇÃO E CORREÇÃO DA BANDA BRILHANTE EM DADOS DE
RADAR METEOROLÓGICO

Dissertação apresentada ao Curso de Pós-Graduação em Métodos Numéricos em Engenharia, Área de Concentração em Programação Matemática, do Departamento de Matemática, Setor de Ciências Exatas e do Departamento de Construção Civil, Setor de Tecnologia, Universidade Federal do Paraná, como parte das exigências para a obtenção do título de Mestre em Ciências.

Orientador: Prof^º. Dr. Sérgio Scheer

Co-Orientador: Dr. Cesar Augustus Assis Beneti

CURITIBA

2014

O48i

Oliveira, Camila de.

Identificação e correção da banda brilhante em dados de radar meteorológico./ Camila de Oliveira. – Curitiba, 2014.
73f. : il. [algumas color.]; 30 cm.

Dissertação (Mestre) - Universidade Federal do Paraná, Setor de Tecnologia, Departamento de Matemática, Setor de Ciências Exatas e do Departamento de Construção Civil, Programa de Pós-graduação em Métodos Numéricos em Engenharia, 2014.

Orientador: Sérgio Scheer. Coorientador: Cesar Augustus Assis Beneti.
Bibliografia: p. 71-73.

1. Radar meteorológico. I. Scheer, Sérgio. II. Beneti, Cesar Augustus Assis. III. Universidade Federal do Paraná. IV. Título.

CDD: 551.6353

TERMO DE APROVAÇÃO

CAMILA DE OLIVEIRA

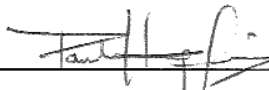
IDENTIFICAÇÃO E CORREÇÃO DA BANDA BRILHANTE EM DADOS DE RADAR METEOROLÓGICO.

Dissertação aprovada como requisito parcial para obtenção do grau de mestre no Programa de Pós-Graduação em Métodos Numéricos em Engenharia, da Universidade Federal do Paraná, pela seguinte banca examinadora:



Prof. Dr. Sérgio Scheer.

Orientador – Membro do PPGMNE/UFPR



Prof. Dr. Paulo Henrique Siqueira.

Membro do PPGMNE/UFPR



Dr. Leonardo Calvetti.

Membro do SIMEPAR.

Curitiba, 06 de junho 2014

AGRADECIMENTOS

Aos meus pais, Amauri e Therezinha, pelo apoio, incentivo e compreensão.

À todos os meus amigos que sempre me apoiaram nos estudos e nas horas de dificuldades.

Ao Prof. Dr. Sérgio Scheer e ao Dr. Cesar Beneti que me orientaram neste trabalho, ajudando em momentos de dúvida.

À todos no SIMEPAR que me ajudaram no desenvolvimento deste trabalho.

RESUMO

A banda brilhante (BB) é uma camada de alta refletividade que ocorre durante o derretimento de neve agregada e cristais de gelo. Essa camada de alta refletividade pode causar erros no cálculo da taxa de precipitação. Por isso, é importante identificar corretamente essa camada nos dados de radar meteorológico para que uma correção possa ser aplicada. Este trabalho utiliza duas técnicas para a identificação desta região. A primeira utiliza os produtos VIL e Z_{max} para classificar os dados em precipitação convectiva e estratiforme, para então, dentro da região estratiforme, limitar a banda brilhante utilizando um cálculo de sua espessura aparente. A segunda, utiliza a metodologia *k-means*, que agrupa os dados em três grupos utilizando variáveis obtidas a partir dos dados de refletividade. Estes três grupos representam a precipitação convectiva, a estratiforme e a região de banda brilhante. Após a identificação da região de banda brilhante é aplicada uma técnica para a correção desses dados utilizando o perfil vertical de refletividade. Os valores de refletividade são então convertidos para taxa de precipitação utilizando uma relação *Z-R* adequada. Em seguida, esses valores são comparados com a precipitação medida por pluviômetros localizados no raio de alcance do radar. As duas técnicas obtiveram melhores resultados quando comparadas com os dados originais. Na comparação entre as técnicas, o *k-means* mostrou, em geral, os melhores resultados.

Palavras-chave: Radar Meteorológico, Banda Brilhante, Perfil Vertical de Refletividade, K-means.

ABSTRACT

The bright band (BB) is a layer of enhanced reflectivity due to melting of aggregated snow and ice crystals. This layer of high reflectivity can cause errors in the calculation of the rain rate. Therefore, it is important to correctly identify this layer in radar observations for a correction to be applied. This current study uses two techniques for the identification of this region. The first uses the products VIL and composite reflectivities to convective-stratiform precipitation segregation, and then, within the stratiform region, limit the bright band using a first-guess apparent BB depth. The other method uses *k-means*, which is used for clustering the data into three clusters, using variables obtained from radar reflectivity. These three clusters represent the convective and stratiform precipitation and bright band region. After identifying the region of bright band, a technique for the correction of these data using the vertical profile of reflectivity is applied. The reflectivity values are then converted to rain rate using an appropriate *Z-R* relationship. Then these values are compared with rainfall measured by rain gauges located within radar range. Both techniques present better results when compared with the original data. However, when comparing the two techniques, *k-means* showed, in general, the best results.

Key-words: Radar, Bright Band, Vertical Profile of Reflectivity, K-means.

LISTA DE FIGURAS

FIGURA 1 – MAPA DA ÁREA DE COBERTURA DO RADAR	16
FIGURA 2 – CÁLCULO DA DISTÂNCIA DO ALVO AO RADAR	17
FIGURA 3 – ESQUEMA DE VARREDURA VOLUMÉTRICA DO RADAR METE- OROLÓGICO	18
FIGURA 4 – ESTRATÉGIA DE COLETA DE DADOS VOLUMÉTRICOS	18
FIGURA 5 – COORDENADAS DOS DADOS DE RADAR	19
FIGURA 6 – VARIAÇÃO DO VOLUME ANALISADO DEVIDO À ABERTURA DO FEIXE	20
FIGURA 7 – EXEMPLO DE 1 M ³ DE FEIXE DE RADAR COM UMA DISTRIBUI- ÇÃO DE GOTAS DE TAMANHOS DIFERENTES	21
FIGURA 8 – PRODUTO PPI	25
FIGURA 9 – PPI DA VARIÁVEL REFLETIVIDADE DO DIA 02/12/2009 ÀS 23:00	25
FIGURA 10– PRODUTO Z_{MAX}	26
FIGURA 11– PRODUTOS Z_{MAX} E VIL DO DIA 02/12/2009 ÀS 23:00	27
FIGURA 12– MODELO CONCEITUAL DE UM CORTE VERTICAL DE UMA LI- NHA DE INSTABILIDADE	31
FIGURA 13– PVR ESQUEMÁTICO PARA DADOS AFETADOS PELA BB	32
FIGURA 14– ILUSTRAÇÃO DO ALGORITMO <i>K-MEANS</i>	41
FIGURA 15– EXEMPLO DO ALGORITMO <i>K-MEANS</i> COM DIFERENTES INI- CIALIZAÇÕES DOS CENTROIDES	42
FIGURA 16– LOCALIZAÇÃO DAS ESTAÇÕES METEOROLÓGICAS DO SIME- PAR E INMET	44
FIGURA 17– ECO DE TERRENO NOS DADOS VOLUMÉTRICOS DE RADAR	45
FIGURA 18– DISTRIBUIÇÃO ANUAL DO PERFIL VERTICAL DE TEMPERA- TURA MÉDIA	46

FIGURA 19– PARÂMETRO D_1	49
FIGURA 20– PARÂMETROS OBTIDOS A PARTIR DO PVR	53
FIGURA 21– PPIS DA ELEVAÇÃO DE 0,5° DO DIA 16/07/2010 ÀS 3:00	56
FIGURA 22– PPIS DA ELEVAÇÃO DE 0,5° DO DIA 22/11/2010 ÀS 5:00	57
FIGURA 23– PPIS DA ELEVAÇÃO DE 0,5° DO DIA 12/04/2011 ÀS 18:00	58
FIGURA 24– TAXA DE PRECIPITAÇÃO (16/07/2010 ÀS 03:00)	61
FIGURA 25– TAXA DE PRECIPITAÇÃO (22/11/2010 ÀS 05:00)	62
FIGURA 26– TAXA DE PRECIPITAÇÃO (12/04/2011 ÀS 18:00)	63
FIGURA 27– VIÉS PARA CADA EVENTO	64
FIGURA 28– REMQ PARA CADA EVENTO	65
FIGURA 29– CLASSIFICAÇÃO DOS HIDROMETEOROS PELOS DOIS MÉTO- DOS (15/10/2010 ÀS 16:00)	67
FIGURA 30– PVR E CORTES VERTICAIS (15/10/2010 ÀS 16:00)	68

LISTA DE TABELAS

TABELA 1	– VIÉS E REMQ PARA O TOTAL DE DADOS OBTIDOS DOS 68 EVENTOS	60
TABELA 2	– VALORES DA TAXA DE PRECIPITAÇÃO PARA 16/07/2010 ÀS 3:00	61
TABELA 3	– VALORES DE VIÉS E REMQ PARA 16/07/2010 ÀS 3:00	61
TABELA 4	– VALORES DA TAXA DE PRECIPITAÇÃO PARA 22/11/2010 ÀS 5:00	62
TABELA 5	– VALORES DE VIÉS E REMQ PARA 22/11/2010 ÀS 5:00	62
TABELA 6	– VALORES DA TAXA DE PRECIPITAÇÃO PARA 12/04/2011 ÀS 18:00	63
TABELA 7	– VALORES DE VIÉS E REMQ PARA 12/04/2011 ÀS 18:00	63

LISTA DE SIGLAS

BB	Banda Brilhante
INMET	Instituto Nacional de Meteorologia
PVR	Perfil Vertical de Refletividade
PPI	<i>Plan Position Indicator</i>
REMQ	Raiz do Erro Médio Quadrático
SIMEPAR	Sistema Meteorológico do Paraná
UTC	<i>Universal Time Coordinated</i>
VIL	<i>Vertically Integrated Liquid-Water</i>

SUMÁRIO

1 INTRODUÇÃO	12
1.1 IMPORTÂNCIA	13
1.2 OBJETIVOS	13
1.3 ESTRUTURA DO TRABALHO	14
2 CONCEITOS METEOROLÓGICOS	15
2.1 RADAR METEOROLÓGICO	15
2.1.1 Coleta de Dados	17
2.1.2 Variáveis Medidas	20
2.1.2.1 Refletividade	21
2.1.2.2 Velocidade Radial	23
2.1.2.3 Largura Espectral	24
2.1.3 Produtos	24
2.1.3.1 PPI	24
2.1.3.2 Z_{max} - Máxima Refletividade da Coluna	25
2.1.3.3 VIL	26
2.2 RELAÇÃO Z-R	27
2.3 PERFIL VERTICAL DE REFLETIVIDADE	28
2.4 PRECIPITAÇÕES CONVECTIVAS E ESTRATIFORMES	29
2.5 BANDA BRILHANTE	31
3 ANÁLISE DE AGRUPAMENTOS	36
3.1 ALGORITMOS <i>HARD CLUSTERING</i>	36
3.1.1 Hierárquicos	37
3.1.2 Não-hierárquicos	37

3.2	ALGORITMOS <i>SOFT CLUSTERING</i>	37
3.3	APLICAÇÕES DA ANÁLISE DE AGRUPAMENTOS	38
3.4	ALGORITMO <i>K-MEANS</i>	39
3.5	ALGORITMO <i>K-MEANS++</i>	42
4	METODOLOGIA	43
4.1	DADOS	43
4.2	CONTROLE DE QUALIDADE DOS DADOS DE REFLETIVIDADE	44
4.3	ISOTERMA DE ZERO GRAU	45
4.4	PYTHON	47
4.5	IDENTIFICAÇÃO DA REGIÃO DE BANDA BRILHANTE	47
4.5.1	Identificação Pela Espessura Aparente	48
4.5.2	Identificação Por <i>K-means</i>	49
4.6	CORREÇÃO DA BANDA BRILHANTE	51
4.6.1	Cálculo do PVR	51
4.6.2	Aplicação da Correção	53
5	RESULTADOS	55
5.1	AVALIAÇÃO DOS MÉTODOS	59
6	CONCLUSÕES	69
6.1	SUGESTÕES PARA TRABALHOS FUTUROS	70
	REFERÊNCIAS	71

1 INTRODUÇÃO

A importância da estimativa confiável de precipitação vem crescendo significativamente nos últimos anos, pois esta é a principal forçante do ciclo hidrológico e assume importância em diversos setores como gerenciamento de águas, agricultura e prevenção de desastres naturais (FILHO; CONDE; ANDRIOLI, 2013).

A precipitação pode ser estimada por métodos diretos ou indiretos. A técnica mais comum de medição direta de precipitação é através de pluviômetros cuja grande limitação é o seu caráter de medida pontual. Para compensar esse caráter pontual, emprega-se rede de pluviômetros para dar um caráter mais extensivo à medida da precipitação (ALMEIDA, 2002). Os processos de métodos indiretos são aqueles em que a estimativa é feita por observação à distância, através de sensoriamento remoto.

Uma das melhores ferramentas de sensoriamento remoto disponíveis atualmente para a elaboração de estimativa de precipitação é o radar meteorológico. Suas principais vantagens são o monitoramento em tempo real de uma ampla área com alta resolução espacial e temporal e a possibilidade de observar a estrutura tridimensional do sistema que deu origem à precipitação.

Porém, as estimativas de chuva com radares meteorológicos podem conter erros provocados por vários fatores. Entre eles está a banda brilhante (BB), termo usado na meteorologia para se referir a uma camada de alta refletividade associada com o derretimento de gelo e que pode superestimar o cálculo da taxa de precipitação. Assim, visando melhorar as estimativas de chuvas a partir dos dados de radar, é necessário que essa região seja identificada corretamente, para que uma correção possa ser aplicada.

1.1 IMPORTÂNCIA

O Sistema Meteorológico do Paraná (SIMEPAR) possui um radar meteorológico Banda-S, Doppler, com polarização simples. Este radar está localizado no município de Teixeira Soares e foi o primeiro radar instalado no Paraná com a finalidade de monitoramento e vigilância do tempo.

Juntamente com o sistema de detecção de descargas atmosféricas e a rede de estações meteorológicas automáticas de superfície, também operados pelo SIMEPAR, as informações disponibilizadas pelo radar meteorológico possibilitam um melhor desempenho na vigilância de eventos severos e previsão a curto prazo da precipitação (BENETI; NOZU; SARAIVA, 1998).

Porém, radares meteorológicos com polarização simples não são capazes de classificar os hidrometeoros, e conseqüentemente, não identificam a banda brilhante corretamente. Por isso, é necessário realizar estudos com os dados obtidos por esses radares para que se possa identificar a banda brilhante e corrigir os valores desta região. Sem uma correta identificação, os elevados valores de refletividade da camada de banda brilhante podem superestimar o cálculo da taxa de precipitação e, como a identificação manual por especialistas é um trabalho custoso, justifica-se a necessidade de utilizar métodos automatizados para a identificação e correção desses dados em tempo real.

1.2 OBJETIVOS

O objetivo principal deste trabalho é fazer a identificação da região de banda brilhante para que uma correção possa ser aplicada. A correta identificação da região afetada é a parte mais importante do trabalho, por isso, duas técnicas serão utilizadas com o objetivo de constatar a que apresenta os melhores resultados. A primeira técnica limita a BB utilizando o cálculo de sua espessura aparente e a segunda faz a identificação utilizando a técnica de análise de agrupamentos *k-means* (k-médias).

Em seguida, será aplicada uma correção nos dados afetados através do uso dos perfis verticais de refletividade da área afetada.

1.3 ESTRUTURA DO TRABALHO

O presente trabalho se inicia com a introdução, na qual é mostrada a importância do trabalho e os seus objetivos. O Capítulo 2 apresenta alguns conceitos sobre radar aplicado à meteorologia. No Capítulo 3 são descritos alguns conceitos sobre a análise de agrupamentos e sobre a técnica *k-means*. Os métodos utilizados para a identificação da banda brilhante e um método para a correção dos dados afetados são descritos no Capítulo 4. O Capítulo 5 apresenta os resultados obtidos, incluindo comparação entre os resultados de cada técnica. E por fim, no Capítulo 6 é feita conclusão do trabalho e sugestões para trabalhos futuros.

2 CONCEITOS METEOROLÓGICOS

O estudo das condições meteorológicas é um objetivo do ramo da ciência denominada meteorologia. Os fenômenos meteorológicos são estudados a partir das observações, experiências e métodos científicos de análise. A observação meteorológica é uma avaliação ou uma medida de um ou vários parâmetros meteorológicos. As observações são sensoriais quando são adquiridas por um observador sem ajuda de instrumentos de medição, e instrumentais, em geral chamadas medições meteorológicas, quando são realizadas com instrumentos meteorológicos (INMET, 2014).

O radar meteorológico é uma das tecnologias mais precisas e utilizadas dentro da meteorologia (PINHEIRO; VAZ; MARTINHAGO, 2005). Ele é capaz de prover informações complexas sobre as dimensões horizontal e vertical da precipitação. Quando utilizado em conjunto com um experimento de medição de taxa de precipitação, permite o levantamento adequado do comportamento local da chuva.

A seguir são apresentados alguns fundamentos básicos sobre radar aplicado à meteorologia, bem como a formação da banda brilhante (BB), cuja identificação e correção é o principal objetivo deste trabalho.

2.1 RADAR METEOROLÓGICO

Radares são dispositivos que permitem detectar objetos a longas distâncias. Um radar meteorológico consiste basicamente de quatro componentes principais: um transmissor para gerar o sinal de alta frequência, uma antena para enviar o sinal e receber o eco de volta do alvo, um receptor para detectar e amplificar o sinal, e algum tipo de sistema de exibição para permitir que se veja o que o radar detectou (RINEHART, 2004).

Os dados utilizados neste trabalho são provenientes do radar Banda-S Doppler, com polarização simples, modelo DWSR-95S, operado pelo SIMEPAR. Este radar está instalado no município de Teixeira Soares (região central do Paraná) e está em operação desde 1998. A cobertura qualitativa deste radar tem raio de 480 km e abrange todo o Paraná, Santa Catarina, parte do sul de São Paulo e norte do Rio Grande do Sul. A área quantitativa (com estimativas de precipitação e vento) é de 240 km de raio, cobrindo todo o centro e leste do Paraná. Estas áreas de cobertura são mostradas na Figura 1.

Os dados do radar do SIMEPAR são obtidos em tempo real para monitoramento e previsão de tempo e armazenados para serem utilizados em pesquisa e desenvolvimento de produtos meteorológicos (NETO, 2008).

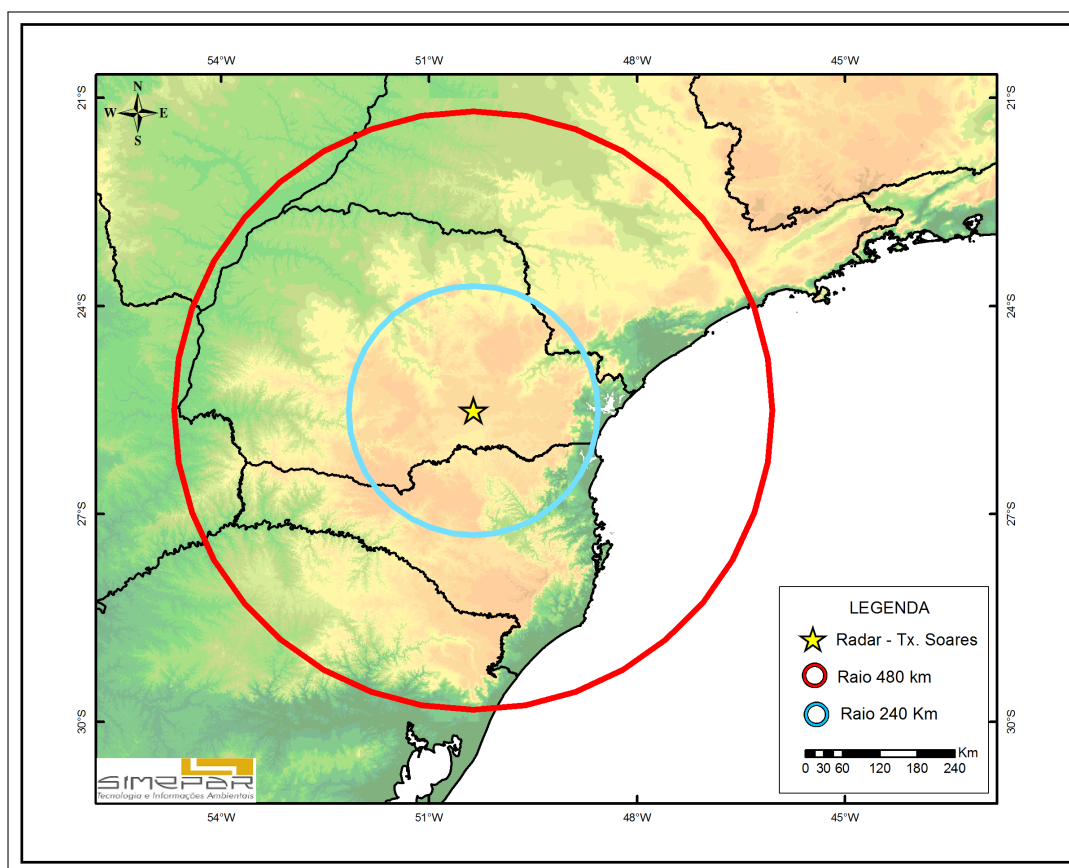


FIGURA 1: Mapa da cobertura qualitativa (480 km de raio) e quantitativa (240 km de raio) do radar meteorológico localizado em Teixeira Soares - PR.

FONTE: SIMEPAR (2014).

2.1.1 COLETA DE DADOS

O radar meteorológico é um sensor ativo que transmite pulsos de energia eletromagnética em intervalos de tempo regulares, concentrados em um feixe de pequena abertura através de uma antena rotativa. Os alvos iluminados pelo feixe do radar podem absorver parte desta energia e irradiá-la em diversas direções. A fração da energia que retorna ao radar é medida pelo receptor e, conhecendo-se o intervalo de tempo t entre a emissão e o retorno do eco, é possível determinar a distância r do alvo (QUEIROZ, 2009), como ilustrado na Figura 2.

Considerando a velocidade de propagação das ondas eletromagnéticas no meio igual a c_m , o intervalo de tempo t é relacionado à distância r pela equação:

$$r = \frac{c_m t}{2}. \quad (1)$$

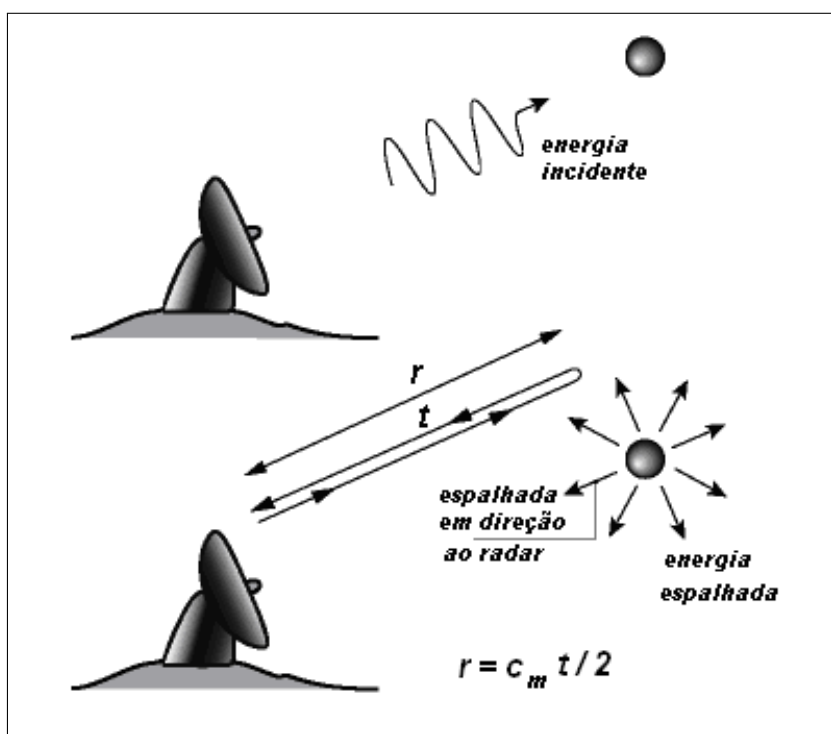


FIGURA 2: Através do tempo de retorno de um feixe emitido é possível calcular a distância do alvo em relação ao radar.

FONTE: Queiroz (2009).

A antena do radar DWSR-95S, de 8,2 m de diâmetro, gera um feixe de ondas eletromagnéticas com 1° de abertura e monitora continuamente a atmosfera com uma sequência pré-programada de varreduras azimutais em 360° . Um volume de varredura corresponde a uma sequência completa de várias varreduras azimutais com diferentes elevações da antena, como mostrado na Figura 3.

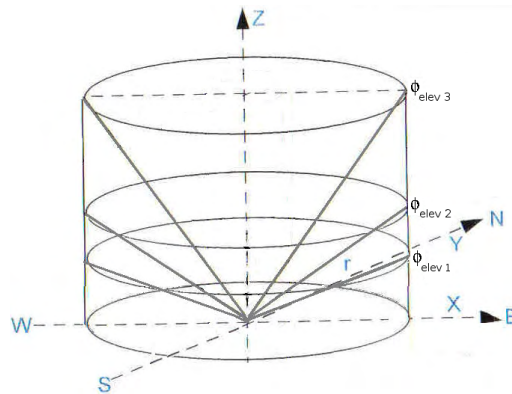


FIGURA 3: Esquema de varredura volumétrica do radar meteorológico.

FONTE: Adaptado de Queiroz (2009).

A estratégia de coleta de dados volumétricos varia de acordo com a ocorrência de precipitação e sua intensidade. A Figura 4 mostra a região iluminada pelo feixe num esquema de varredura com 14 elevações da antena.

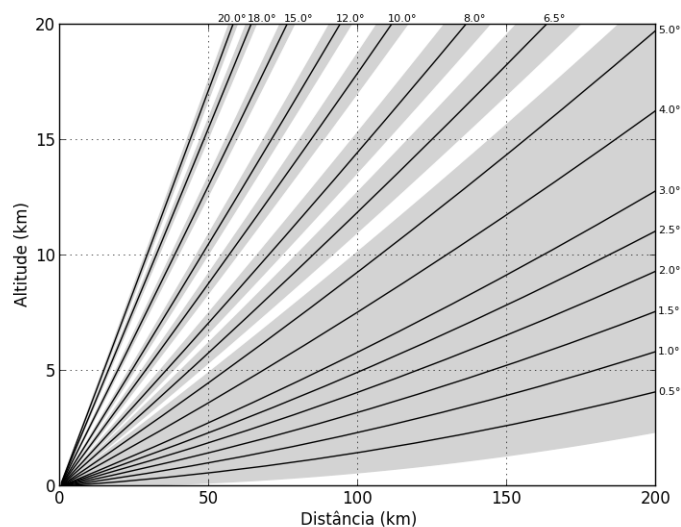


FIGURA 4: Estratégia de coleta de dados volumétricos.

FONTE: A autora (2014).

Os dados coletados pelo radar possuem coordenadas esféricas com três graus de liberdade (r, θ, ϕ) , onde r é a distância do alvo até o radar (*slant range*), θ é o ângulo de azimute em relação ao norte e, ϕ é o ângulo de elevação da antena em relação ao horizonte. Essas três informações determinam a localização do alvo no espaço e são mostradas na Figura 5.

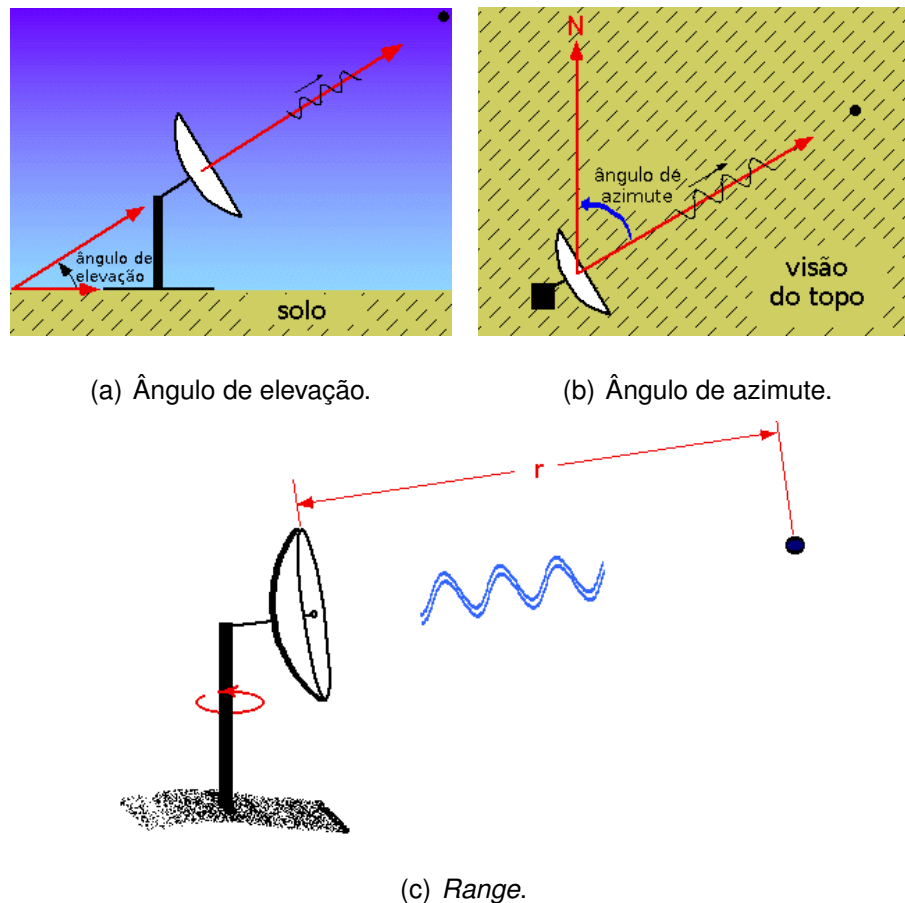


FIGURA 5: Coordenadas dos dados de radar.

FONTE: Adaptado de Illinois (2013).

Devido à abertura do feixe (*beam width*), o feixe apresenta uma forma de cone, onde a metade da potência transmitida pelo radar está concentrada e, quanto maior a distância do radar, maior é o volume analisado, como pode ser observado na Figura 6. Quando o sinal refletido retorna, o software do radar se encarrega de separá-lo em intervalos com comprimento constante na direção radial. Cada unidade de volume criado por essa separação é denominado *bin* e recebe o valor da média ponderada dos

alvos detectados naquela amostra. Assim, dois ou mais alvos que couberem dentro do volume iluminado serão tomados como um alvo só, resultando em um único eco.

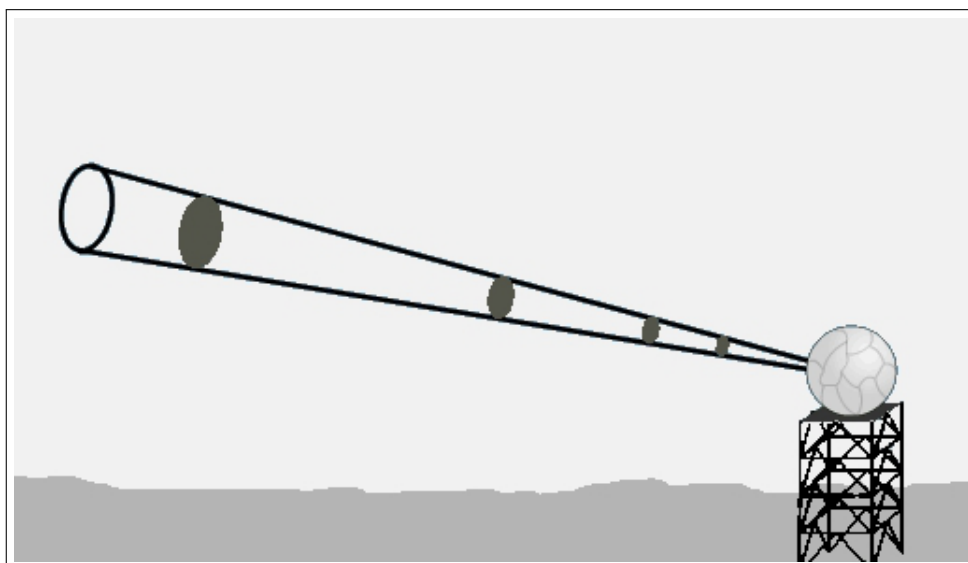


FIGURA 6: Variação do volume analisado devido à abertura do feixe.

FONTE: MeteoPT (2013).

A altitude h de um *bin* é representada pela altitude central do feixe e pode ser calculada através da equação (DOVIAK; ZRNIC, 1993):

$$h = (r^2 + a_e^2 + 2ra_e \sin\phi)^{1/2} - a_e, \quad (2)$$

onde r é o *range*, ϕ é o ângulo de elevação e a_e é chamado de raio efetivo da Terra e corresponde a 4/3 do raio da Terra.

2.1.2 VARIÁVEIS MEDIDAS

O sistema de processamento dos sinais do DWSR-95S gera três variáveis (ou momentos) básicas em meteorologia com radar: refletividade, velocidade radial média e largura espectral.

2.1.2.1 REFLETIVIDADE

A refletividade é um parâmetro meteorológico determinado pelo número e tamanho das partículas presentes no volume de amostragem (RINEHART, 2004). O fator de refletividade do radar pode ser expresso por:

$$z = \sum_{vol} D^6, \quad (3)$$

onde D é o diâmetro da gota e o somatório é feito sobre o número total de gotas de tamanhos variados dentro de uma unidade de volume de feixe (*bin*). Sua unidade é expressa em mm^6m^{-3} .

A Figura 7 mostra um exemplo de uma seção de 1 m^3 do feixe do radar que contém uma distribuição de gotas de tamanhos diferentes.

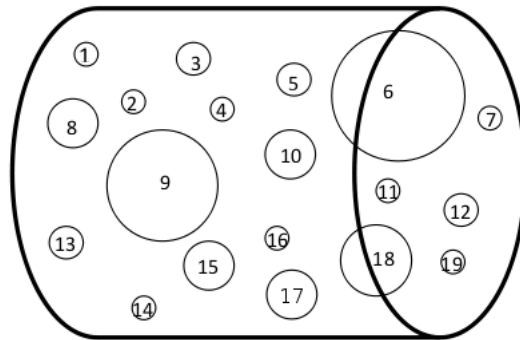


FIGURA 7: Exemplo de 1 m^3 de feixe de radar com uma distribuição de gotas de tamanhos diferentes.

FONTE: Santos (2014).

Assim, o somatório para o fator de refletividade do radar z pode ser expandido da seguinte forma:

$$z = D_1^6 + D_2^6 + \dots + D_{18}^6 + D_{19}^6 \quad (4)$$

Nota-se que z é uma propriedade inerente da distribuição do tamanho da gota e não é dependente do radar. Outro ponto importante é que z é proporcional à sexta potência dos tamanhos da gota, logo, gotas maiores causam grandes valores de refletividade quando comparadas com gotas pequenas (SANTOS, 2014).

Contudo, a distribuição do tamanho da gota no volume medido é desconhecida. Assim, calcula-se o fator de refletividade z pela energia retornada ao radar (p_r). Esta energia recebida está relacionada com o fator de refletividade, entre outros parâmetros do radar pela equação 5 (RINEHART, 2004), denominada como equação do radar.

$$p_r = \frac{\pi^3 p_t g^2 \theta_0 \varphi_0 l |K|^2 z}{1024 \ln(2) \lambda^2 r^2}, \quad (5)$$

onde:

p_t = energia transmitida pelo radar (watts);

g = ganho da antena (razão entre a potência que é recebida em um ponto específico no espaço a partir de uma antena isotrópica);

θ_0 = largura do feixe horizontal (radianos);

φ_0 = altura do feixe vertical (radianos);

l = largura de pulso (metros);

λ = comprimento de onda do pulso do radar (metros);

$|K|^2$ = constante dielétrica para hidrometeoros. Normalmente com valor de 0,93 para água líquida e 0,197 para gelo;

r = distância do alvo (*slant range*);

z = fator de refletividade do radar.

Pode-se simplificar consideravelmente a equação do radar, agrupando os parâmetros associados a um radar específico como uma constante c_1 . Já que o interesse principal é em hidrometeoros líquidos, assume-se o valor de 0,93 para a constante dielétrica $|K|^2$. Tem-se assim, a equação do radar simplificada:

$$p_r = \frac{c_1 z}{r^2} \quad (6)$$

A potência recebida pelo radar é proporcional ao fator de refletividade e inversamente proporcional à distância. Quanto mais intensa for a tempestade, maior é o valor de refletividade e maior será a energia recebida pelo radar.

A variação da distância também é importante. Quanto mais distante a tempestade

está do radar, menor é a potência retornada. Duas tempestades de igual intensidade retornarão a mesma energia apenas se estiverem à mesma distância.

Rearranjando a equação 6, tem-se a equação para o cálculo do fator de refletividade do radar:

$$z = \frac{1}{c_1} p_r r^2 = c_2 p_r r^2 \quad (7)$$

A refletividade pode variar desde valores próximos de $0,001 \text{ mm}^6\text{m}^{-3}$ (névoa úmida) até $36.000.000 \text{ mm}^6\text{m}^{-3}$ (granizo grande ~ 5 a 10 cm) (QUEIROZ, 2009). Por isso é conveniente utilizar sua representação em uma escala logarítmica:

$$Z = 10 \log \left(\frac{z}{1 \text{ mm}^6\text{m}^{-3}} \right), \quad (8)$$

onde Z é o valor de refletividade em escala logarítmica, medida em dBZ, e z é o valor de refletividade em escala linear (mm^6m^{-3}).

2.1.2.2 VELOCIDADE RADIAL

O radar mede a velocidade do vento a partir da velocidade dos hidrometeoros. Esta variável facilita prever para onde o evento se deslocará.

O radar utiliza o efeito doppler para mensurar a velocidade radial do vento. O efeito Doppler consiste numa mudança de frequência na onda causada pelo movimento da fonte de onda. Por exemplo, quando uma ambulância se aproxima, o ruído feito por ela aumenta, enquanto que quando se afasta, o ruído diminui. O mesmo princípio se aplica às ondas eletromagnéticas do radar: a partir da mudança da frequência na onda, pode-se determinar se o alvo está se afastando ou se aproximando do radar e qual a sua velocidade de deslocamento (SANTOS, 2011).

O vento radial é mostrado em m/s com valores positivos para os alvos que se afastam e valores negativos para os alvos que se aproximam do radar.

2.1.2.3 LARGURA ESPECTRAL

A Largura Espectral mede o desvio padrão das medidas obtidas em cada amostra (*bin*) e, meteorologicamente, indica a turbulência nesse volume conforme a variância dos alvos presentes. Esta variável também é medida em m/s.

2.1.3 PRODUTOS

Os dados obtidos pelo radar chegam em formato numérico e, para uma melhor percepção humana, são convertidos para imagens. Essas imagens (produtos) podem ser visualizadas de diferentes modos, e alguns deles são apresentados a seguir.

2.1.3.1 PPI

O PPI (*Plan Position Indicator*) fornece a projeção num plano horizontal, obtido através de uma varredura em azimuth em coordenadas polares, para um ângulo de elevação constante. Esta é a forma básica de visualização das variáveis medidas pelo radar.

Deve-se observar que, conforme a distância em relação ao radar aumenta, a altitude do feixe com relação ao solo também aumenta. A Figura 8 apresenta a varredura do radar em um determinado ângulo de elevação com plano fixo e ilustra também como a altitude do feixe aumenta com o *range*.

Os dados em coordenadas esféricas são transformados para coordenadas cartesianas e então, são plotados com o centro da imagem indicando a localização do radar. A Figura 9 mostra PPIs da variável refletividade obtidos da varredura realizada pelo radar no dia 02/12/2009 às 23:00. Na Figura 9(a) o PPI é obtido da elevação de $0,5^\circ$ e na Figura 9(b) tem-se o PPI da elevação de $2,5^\circ$. As cores quentes representam valores de refletividade mais elevados, enquanto que cores frias representam valores mais baixos.

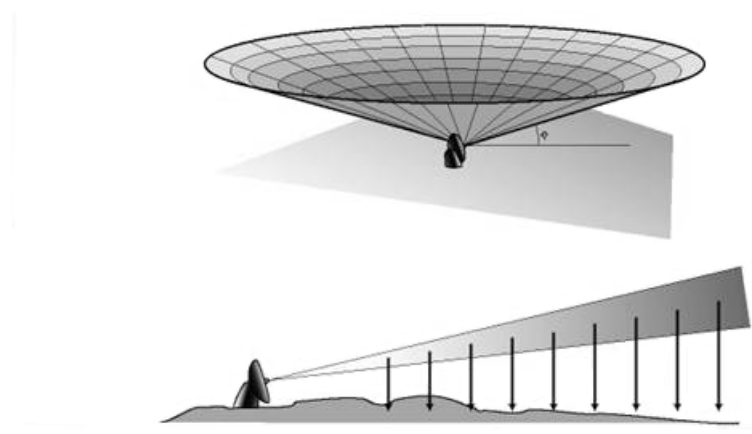


FIGURA 8: Produto PPI.

FONTE: MeteoPT (2013).

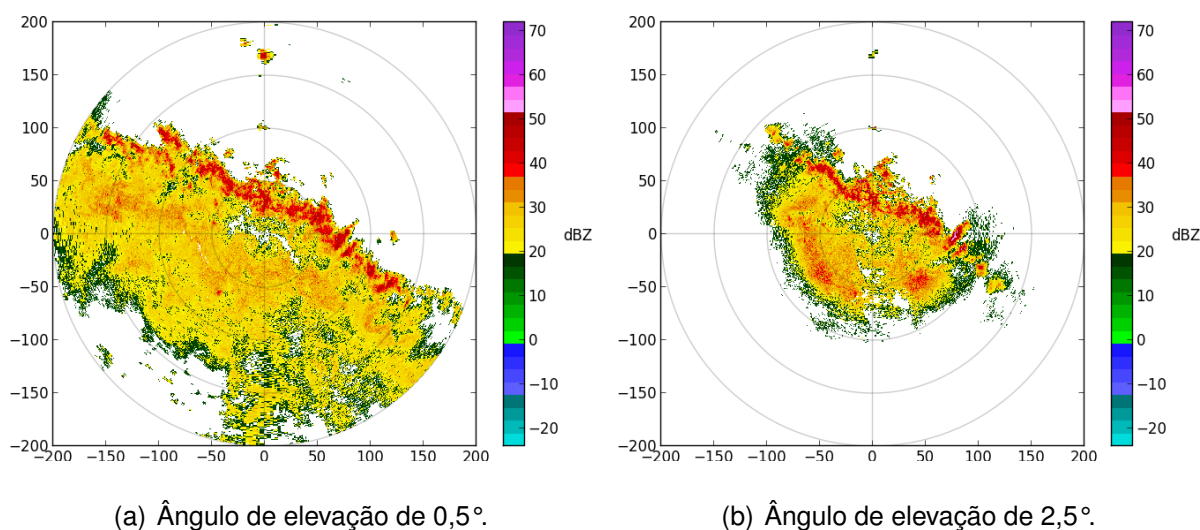
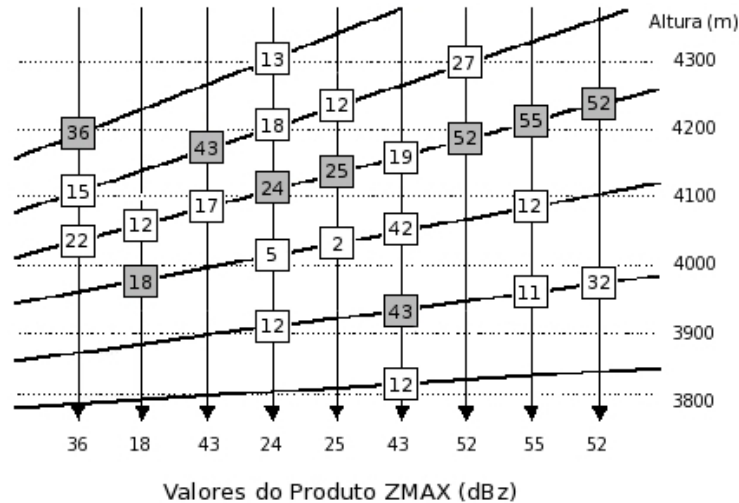


FIGURA 9: PPI da variável refletividade do dia 02/12/2009 às 23:00 com um raio de alcance de 200 km. Em 8(a) tem-se um ângulo de elevação de 0,5° e em 8(b) um ângulo de elevação de 2,5°. O centro da imagem representa a localização do radar.

FONTE: A autora (2014).

2.1.3.2 Z_{MAX} - MÁXIMA REFLETIVIDADE DA COLUNA

O Z_{max} indica a refletividade máxima encontrada em cada coluna. Ou seja, para cada posição (r, θ) dos dados do radar, encontra-se o maior valor de refletividade em todas as elevações do volume de varredura, como ilustrado na Figura 10. Um exemplo de imagem do produto Z_{max} é mostrado na Figura 11(a).

FIGURA 10: Produto Z_{max} .

FONTE: MeteoPT (2013).

2.1.3.3 VIL

O termo VIL, do inglês *Vertically Integrated Liquid-Water*, representa o conteúdo de água líquida integrado verticalmente, ou seja, os dados de refletividade são convertidos em estimativas de conteúdo de água líquida com base em estudos teóricos de distribuição de tamanho de gota e estudos empíricos do fator de refletividade. Esta técnica apresenta as características tridimensionais de um sistema de tempestades em uma exibição bidimensional (GREENE; CLARK, 1972), fornecendo um meio para a identificação e possível previsão de tempestades severas.

Uma forma de cálculo do VIL é descrita em Amburn e Wolf (1997), e é dada pela equação:

$$VIL = 3,44 \times 10^{-6} \sum \left(\frac{z_i + z_{i+1}}{2} \right)^{4/7} \Delta h \quad (9)$$

onde z_i e z_{i+1} são os valores de refletividade (em mm^6m^{-3}) no limite inferior e superior, respectivamente, de uma camada de amostra e Δh é a diferença de altura (em metros) entre os limites dessa camada.

O VIL é calculado a partir de um único volume de varredura e é expresso em kg/m^2 . Um exemplo de imagem do produto VIL é mostrado na Figura 11(b).

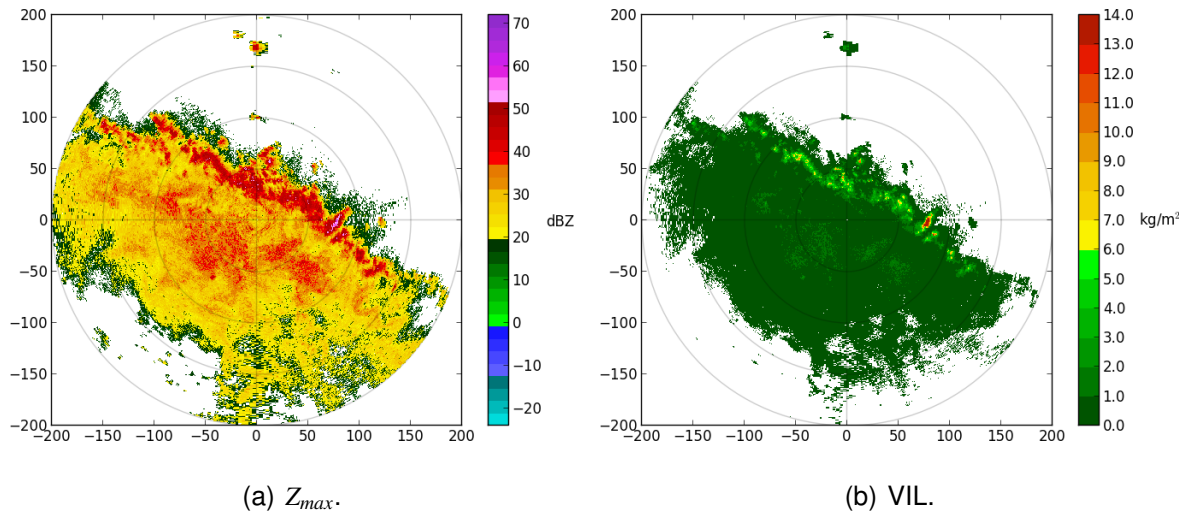


FIGURA 11: Produtos Z_{max} e VIL obtidos do volume de varredura do dia 02/12/2009 às 23:00 para um raio de 200 km.

FONTE: A autora (2014).

2.2 RELAÇÃO Z-R

O radar não mede diretamente a chuva, ele recebe uma determinada densidade de potência refletida pelos alvos de chuva denominado refletividade. Esta refletividade possui uma relação física com o espectro de gotas observado. A partir deste espectro pode-se determinar uma relação entre a refletividade do radar e a taxa de precipitação correspondente (QUEIROZ, 2009).

A taxa de precipitação R é definida como o volume equivalente de água líquida que cruza uma unidade de área de seção horizontal por unidade de tempo (ALMEIDA, 2002).

Uma relação empírica pode ser obtida plotando-se gráficos de taxa de precipitação que chega no solo contra a refletividade calculada pelo radar. Sua forma mais conhecida e utilizada é

$$z = aR^b \quad (11)$$

onde z é a refletividade em mm^6m^{-3} , R é a taxa de precipitação em mm/h , e a e b são constantes empíricas, determinadas por meio da distribuição das partículas.

Os valores dessas constantes variam de acordo com a região do globo e tipos de precipitação (DAMIAN, 2011).

A relação mais frequentemente usada na literatura é a proposta por Marshall e Palmer (1948), que utiliza como constantes empíricas os valores 200 e 1,6 para a e b , respectivamente.

Porém, Santos (2014) após estudos com dados do SIMEPAR para o radar DWSR-95S de Teixeira Soares, propôs uma nova relação que se mostrou melhor para a estimativa de precipitação desta região. Esta nova relação Z - R é utilizada para todos os dados sem distinção dos tipos de precipitação e possui o valor 52 para a constante a e 2,7 para b :

$$z = 52 R^{2,7}. \quad (12)$$

2.3 PERFIL VERTICAL DE REFLETIVIDADE

A refletividade $Z_t(r, \theta, \phi)$ depende da localização geográfica do ponto (x) e de sua altitude h associada com (r, ϕ) e pode ser aproximada por:

$$Z_t(r, \theta, \phi) \approx Z_t(x, h) \quad (13)$$

Assumindo que esta função pode ser decomposta em duas funções independentes, tem-se:

$$Z_t(x, h) = Z_t(x, 0)z_t(h) \quad (14)$$

onde $Z_t(x, 0)$ é o valor de refletividade ao nível do solo e $z_t(h)$ é chamado de perfil vertical de refletividade, ou PVR verdadeiro.

O perfil calculado diretamente a partir dos dados obtidos pelo radar é chamado de PVR aparente (z_a). Este perfil difere do PVR verdadeiro pois inclui os efeitos de ampliação do feixe, que dependem das características do radar como o comprimento de onda emitida (λ), abertura do feixe (ϕ_0) e diâmetro da antena (D). O PVR aparente

pode ser relacionado com o PVR verdadeiro a partir da seguinte equação:

$$Z_a(x, h) = \int_{H^-}^{H^+} Z_t(x, y) f^4(y) dy = Z_t(x, 0) z_a(h) \quad (15)$$

onde

$$z_a(h) = \int_{H^-}^{H^+} z_t(y) f^4(y) dy. \quad (16)$$

H^- e H^+ representam os limites inferior e superior, respectivamente, do feixe emitido e f^4 representa o quadrado da distribuição de energia do feixe com a altitude h , a qual pode ser aproximada pela função de Bessel de segunda ordem (DOVIAK; ZRNIC, 1993):

$$f^2(\alpha) = \left\{ \frac{8J_2[(\pi D \sin \alpha)/\lambda]}{[(\pi D \sin \alpha)/\lambda]^2} \right\}^2 \quad (17)$$

Tomando a equação 15 em escala logarítmica e rearranjando os termos, tem-se o valor da refletividade ao nível do solo:

$$dBZ_t(x, 0) = dBZ_a(x, h) - dBz_a(h) \quad (18)$$

Além de desempenhar um papel importante na estimativa de precipitação derivada a partir de medições de radar, através deste perfil pode-se, por exemplo, identificar a precipitação como sendo apenas uma garoa ou uma tempestade. Ele pode indicar também a presença de gelo e da banda brilhante, a qual afeta consideravelmente o cálculo da taxa de precipitação na superfície.

2.4 PRECIPITAÇÕES CONVECTIVAS E ESTRATIFORMES

Os modos básicos de precipitação são convectiva e estratiforme e a maior parte dos sistemas de precipitação pode ser decomposta nesses modos, embora a distinção não possa ser sempre bem definida.

A classificação de precipitação em regimes convectivo e estratiforme é útil numa variedade de aplicações em meteorologia, pois as propriedades microfísicas e magnitude dos movimentos verticais da chuva associadas com os dois regimes são diferen-

tes.

Sistemas convectivos são associados com fortes campos verticais de velocidade, cobertura de área pequena e alta intensidade de chuvas, enquanto que os sistemas estratiformes têm fracos campos verticais de velocidade, maior homogeneidade horizontal e menor intensidade de chuvas. Embora a taxa de precipitação na área estratiforme seja muito mais fraca do que a de células convectivas, a chuva estratiforme normalmente cobre maiores áreas e sempre contribui com uma parcela significativa do volume de chuva, mesmo para grandes sistemas convectivos (ANAGNOSTOU, 2004).

A precipitação estratiforme existe quando o movimento vertical no interior da nuvem satisfaz a condição:

$$|w| \ll |v_t| \quad (19)$$

onde w é a velocidade do vento e v_t é a velocidade terminal das partículas de gelo. Portanto, as partículas de gelo nas regiões mais altas devem precipitar enquanto crescem, pois não são sustentadas ou carregadas para cima. Nos processos convectivos, os movimentos verticais são da ordem de 1-10 m/s ou mais, igualando ou excedendo as velocidades terminais dos cristais de gelo (ALBRECHT; DIAS, 2004).

A Figura 12 mostra um modelo conceitual de um corte transversal de uma linha de instabilidade indicando o campo de refletividade do radar. Os hidrometeoros da região estratiforme passam pela isoterma de zero grau (nível de congelamento), derretem e formam uma região horizontal de altas refletividades, conhecida como banda brilhante. Esta camada é detectada pelo radar e pode provocar erros no cálculo da taxa de precipitação. Na seção seguinte é apresentada mais detalhadamente a formação dessa camada.

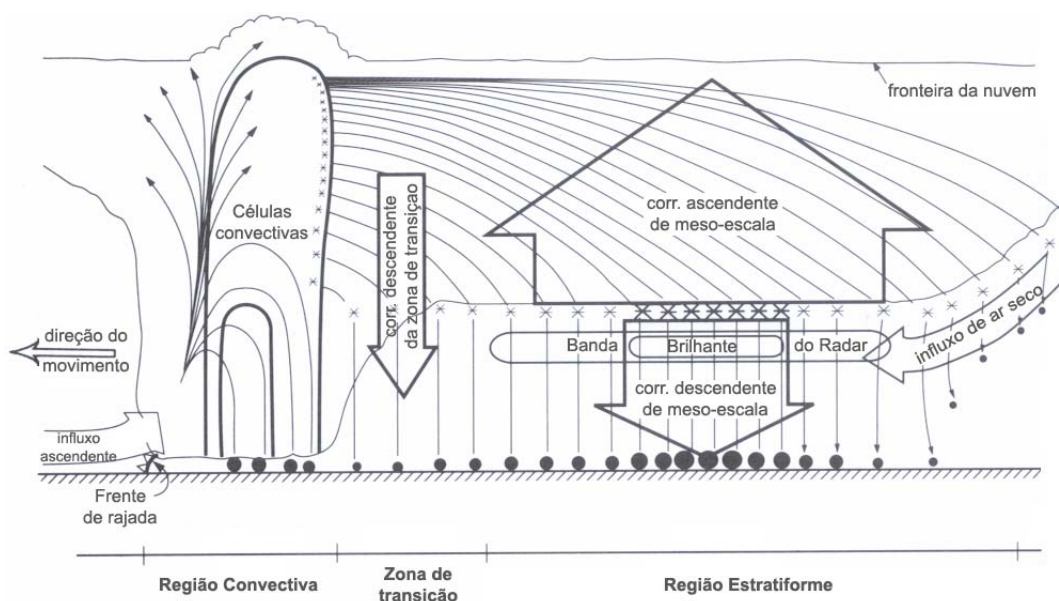


FIGURA 12: Modelo conceitual da trajetória da precipitação e movimentos verticais através de um corte vertical de uma linha de instabilidade. Asteriscos representam os hidrometeoros e os círculos cheios a intensidade da precipitação.

FONTE: Albrecht e Dias (2004).

2.5 BANDA BRILHANTE

A banda brilhante é uma camada de elevada refletividade formada devido ao derretimento de gelo agregado.

Acima da isoterma de zero grau, o gelo cai com uma velocidade terminal relativamente mais lenta e logo que o gelo atinge essa isoterma entra em processo de fusão. Uma vez que o gelo está em temperaturas ligeiramente acima da temperatura de congelamento, começará a derreter de fora para dentro. Isto significa que as extremidades do gelo derreterão primeiro, formando um revestimento de água, enquanto ainda permanece moderadamente grande e de forma irregular. Assim, o radar detecta este gelo como se fosse uma grande gota de água, podendo aumentar significativamente o valor da refletividade lida pelo radar.

Como a neve revestida de água continua a cair e derreter, suas reduções de tamanho e sua velocidade terminal aumentam. A consequência do primeiro efeito é que a refletividade diminui um pouco, dependendo da mudança de diâmetro eficaz entre a

neve e a gota de água. O resultado do segundo efeito é que as gotas que caem abaixo da isoterma se movem mais rapidamente, reduzindo a densidade e concentração dos flocos de gelo, o que diminui ainda mais a refletividade nesta região.

Este efeito na variação do valor da refletividade pode ser observado na Figura 13, onde é apresentado um modelo conceitual para o PVR de dados afetados pela banda brilhante. Este modelo de PVR é assumido para representar as condições atmosféricas através de uma certa região da área do radar.

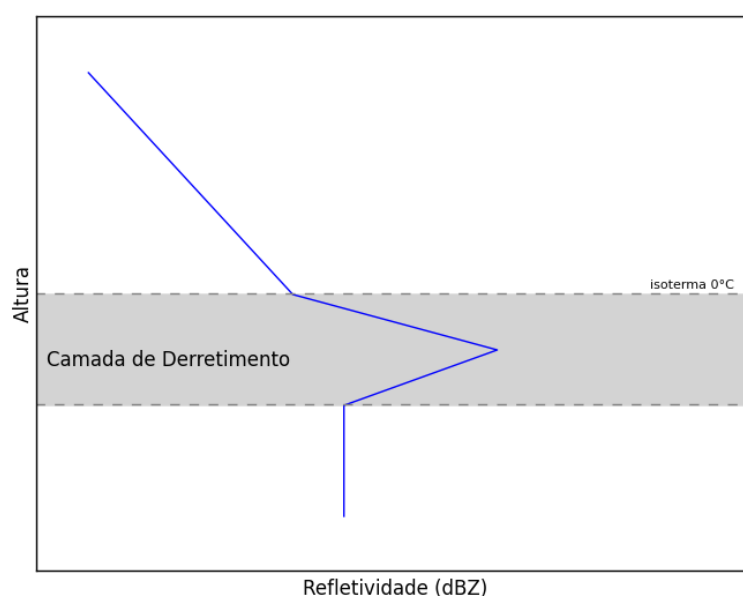


FIGURA 13: PVR esquemático para dados afetados pela BB.

FONTE: A autora (2014).

Essa camada de banda brilhante pode causar significativos erros no cálculo da taxa de precipitação se uma correção não for aplicada, erros esses que acontecem em consequência do perfil vertical de refletividade não uniforme. Como a banda brilhante ocorre principalmente durante a precipitação estratiforme (na precipitação convectiva as mesmas físicas se aplicam, mas a transição entre gelo e chuva muitas vezes é tão caótica que não são detectadas), esta delimitação pode ser feita através do PVR calculado a partir dos dados com esta classificação de precipitação. Este PVR fornece alguns parâmetros que podem ajudar na delimitação desta camada.

Vários métodos para detectar automaticamente esta camada de banda brilhante foram propostos. Fabry e Zawadzki (1995) definiram o topo da BB (BB_{topo}) como a altura em que a curvatura do PVR é máxima e, utilizou também, o mesmo critério para definir a altitude da base (BB_{base}). Outras técnicas para a identificação e correção da banda brilhante a partir de perfis verticais de refletividade podem ser encontrados em Smith (1986), Smyth e Illingworth (1998), Sanchez-Diezma, Zawadzki e Sempere-Torres (2000), Gourley e Calvert (2003), Mittermaier e Illingworth (2003), Zhang, Langston e Howard (2008), Zhang e Qi (2010), entre outros.

Segundo Zhang e Qi (2010), os métodos de identificação e correção da BB podem ser classificados em quatro grupos de acordo com os PVRs utilizados: 1) PVR climatológico; 2) PVR médio calculado a partir de dados volumétricos obtidos próximos ao radar e, em seguida, aplicado para correções nas regiões mais distantes; 3) PVRs locais recuperados; e 4) modelo conceitual ou PVRs parametrizados.

O PVR climatológico pode ser obtido através de estudos ao longo do tempo e espaço. Este perfil não reflete a estrutura vertical da refletividade em tempo real, não detectando assim, eventos como a presença de banda brilhante. Em muitos estudos eles são utilizados como PVR padrão quando os PVRs em tempo real não estão disponíveis.

O PVR médio é calculado em tempo real, combinando várias altitudes de medição de um único volume de varredura ou de vários volumes em um curto intervalo de tempo, podendo assim, ser utilizado para extrapolar dados de radar ao nível do solo. Ele é calculado com os dados de um raio próximo ao radar, pois esses dados são provenientes de feixes razoavelmente livres de bloqueios e não demasiadamente elevados em relação ao nível do solo. Este perfil é então utilizado para corrigir os dados ao longo de todo o domínio do radar. Se não é registrada chuva próximo ao radar, o perfil climatológico é usado no lugar. Geralmente, admite-se que este PVR médio possui a mesma estrutura ao longo de toda a área do radar, negligenciando assim, as variações espaciais dos perfis verticais de precipitações.

Os perfis locais recuperados levam em consideração as variações no tempo e espaço. O perfil fornecido pelos dados de um volume de varredura é chamado de perfil aparente e, devido ao efeito de suavização do feixe, ele difere do perfil verdadeiro. Quando a região de interesse está longe do radar (normalmente mais que 50 km), esta suavização faz o perfil aparente inadequado para corrigir os dados de radar. O perfil verdadeiro é então recuperado utilizando um método inverso, o qual é descrito em Andrieu e Creutin (1995).

Vignal et al. (2000) utilizou esses três tipos de PVR para corrigir a estimativa de precipitação. Seus resultados sugerem que mesmo uma estimativa grosseira do perfil, como por exemplo o PVR climatológico, permite reduzir a diferença do cálculo da taxa de precipitação a partir da relação $Z-R$ comparadas com os dados observados. Porém, os perfis verticais locais foram os que apresentaram melhores resultados para a correção.

No entanto, a abordagem por PVR local é computacionalmente dispendiosa e não é facilmente implementada para aplicações operacionais. Assim, uma abordagem alternativa é utilizar modelos conceituais de PVRs com alguns parâmetros definidos a partir da estrutura da precipitação.

Zhang e Qi (2010) utilizaram uma abordagem que combina o PVR médio obtido de um volume de varredura com um modelo conceitual para desenvolver uma técnica que corrige automaticamente erros causados pelo efeito da banda brilhante em dados em tempo real. Esta técnica é baseada nas seguintes considerações:

1. O PVR médio é utilizado ao invés do PVR local recuperado devido à sua simplicidade e eficiência computacional.
2. Os PVRs médios são calculados a partir de um único volume de varredura ao invés de uma média de várias varreduras em um curto intervalo de tempo. Isto simplifica os cálculos, evitando a necessidade de associações de tempo e diminui erros devido ao deslocamento da tempestade.

3. Um modelo conceitual de PVR é utilizado com um pequeno número de parâmetros que podem ser derivados a partir de informações meteorológicas e a partir dos dados volumétricos do radar. Além da simplicidade, o modelo idealizado pode ser usado para corrigir a precipitação na região de banda brilhante quando ela está perto do solo. Nesta situação, a refletividade abaixo da BB não pode ser obtida a partir de observações de radar e abordagens de correções tradicionais podem não funcionar corretamente.
4. Os parâmetros no modelo conceitual de PVR são ajustados dinamicamente utilizando observações de radar em tempo real, bem como dados ambientais. Isso pode ajudar a obter PVRs mais representativos do que modelos idealizados com parâmetros pré-definidos. Esta última abordagem pode resultar em descontinuidades nos campos de radar corrigidos quando parâmetros pré-definidos não representam a distribuição da BB em tempo real.

A técnica, descrita em Zhang e Qi (2010), para a correção da banda brilhante utilizando PVR, será aplicada nos dados do radar meteorológico do SIMEPAR e será detalhada na seção 4.6.

Porém, antes da correção, é necessário fazer uma correta identificação da região afetada. Este passo é muito importante, pois se houver uma região classificada como BB maior do que a real, correções poderão ser aplicadas em lugares que realmente apresentam chuva intensa. Por outro lado, se a região classificada for menor, o cálculo da taxa de precipitação continuará sendo superestimado.

Por isso, serão utilizadas duas técnicas para a identificação da BB, a fim de se determinar a que apresenta melhores resultados após a aplicação da correção. A primeira técnica utiliza os produtos descritos neste capítulo. A segunda, utiliza a técnica de agrupamentos *k-means*, e sua descrição, juntamente com a teoria de análise de agrupamentos, serão apresentadas no próximo capítulo.

3 ANÁLISE DE AGRUPAMENTOS

A análise de agrupamentos, também chamada de análise de *clusters*, é uma sub-área da Análise Multivariada, a qual basea-se em técnicas que analisam simultaneamente múltiplas medidas de objetos sob análise. Seu propósito consiste em separar objetos em grupos, baseando-se nas características que estes objetos possuem.

Ou seja, têm o objetivo de agrupar um conjunto de dados em um determinado número de grupos de modo que o grau de associação entre dados de um mesmo grupo seja alto e entre dados de grupos diferentes seja baixo. A análise de agrupamentos possui como característica a capacidade de identificar diferentes padrões num conjunto de dados, tornando possível a divisão destes dados em grupos similares.

Os agrupamentos podem ser divididos em dois tipos: *hard clustering* onde cada dado pertence exclusivamente a um grupo e *soft clustering* onde cada dado está associado a todos os grupos por um grau de pertinência.

3.1 ALGORITMOS *HARD CLUSTERING*

Dado um conjunto de N elementos $X = \{x_1, x_2, \dots, x_N\}$, o problema de agrupamento *hard clustering* consiste na obtenção de um conjunto com k grupos $C = \{C_1, C_2, \dots, C_k\}$, tal que os elementos contidos em C_i são agrupados em função de sua similaridade ou dissimilaridade (DAMIAN, 2011). Os grupos C_i devem satisfazer as seguintes condições (THEODORIDIS; KOUTROUMBAS, 2009):

$$\begin{cases} C_i \neq \emptyset, & i = 1, \dots, k \\ \bigcup_{i=1}^k C_i = X \\ C_i \cap C_j = \emptyset, & i, j = 1, \dots, k, \quad i \neq j \end{cases} \quad (20)$$

A partir destas condições, observa-se que cada dado do conjunto deve pertencer

a um grupo e, exclusivamente a este grupo.

As técnicas que realizam *hard clustering* podem ser classificadas em duas principais categorias: hierárquicas e não-hierárquicas.

3.1.1 HIERÁRQUICOS

Os precedimentos hierárquicos envolvem uma série de $N - 1$ decisões de agrupamentos (sendo N o número de observações) que combinam observações em uma estrutura de hierarquia ou do tipo árvore (DAMIAN, 2011).

Dois tipos básicos de métodos hierárquicos podem ser encontrados: Aglomerativos e Divisivos.

- Aglomerativos: inicialmente, cada objeto é considerado um agrupamento. Os agrupamentos mais similares vão se unindo e formando novos agrupamentos até que exista apenas um agrupamento principal. Exemplo: AGNES (*Agglomerative Nesting*).
- Divisivos: todos os objetos iniciam em um único agrupamento e sucessivamente vão sofrendo divisões até que cada objeto represente um agrupamento. Exemplo: DIANA (*Divisive Analysis*).

3.1.2 NÃO-HIERÁRQUICOS

Diferentemente dos métodos hierárquicos, os métodos não-hierárquicos não envolvem o processo de construção em árvore. Em vez disso, designam objetos a agrupamentos assim que o número de agregados a serem formados tenha sido especificado (DAMIAN, 2011). Ou seja, é uma técnica usada quando se deseja formar k grupos de itens ou objetos. Exemplos: *k-means*, *k-medoid*.

3.2 ALGORITMOS *SOFT CLUSTERING*

Algoritmos *soft clustering* associam um dado a todos os grupos através da variação do grau de pertinência deste dado em cada grupo.

A vantagem destes algoritmos é que eles podem representar melhor a incerteza e este fato é importante, por exemplo, para mostrar que um indivíduo não é um típico indivíduo de nenhuma das classes, mas tem similaridade em maior ou menor grau com mais de uma classe (JUNIOR, 2006).

Dentre as técnicas que realizam *soft clustering*, a mais conhecida é a *Fuzzy c-means*.

3.3 APLICAÇÕES DA ANÁLISE DE AGRUPAMENTOS

A análise de agrupamento é uma ferramenta útil que pode ser utilizada em inúmeras variedades de aplicações. Como por exemplo, pode-se citar (THEODORIDIS; KOUTROUMBAS, 2009):

- Redução de dados: para uma grande quantidade de dados a análise de agrupamentos pode ser usada a fim de agrupar dados em um número sensato de grupos e processar cada grupo como uma única entidade.
- Geração de hipóteses: aplica-se a análise de agrupamentos para um conjunto de dados, a fim de inferir algumas hipóteses sobre a natureza dos dados. Assim, o agrupamento é usado aqui como um veículo para sugerir hipóteses. Estas hipóteses devem, então, ser verificadas utilizando outros conjuntos de dados.
- Testes de hipóteses: a análise de agrupamentos pode ser usada para verificar a validade de uma hipótese específica.
- Previsão baseada em grupos: para o conjunto de dados disponível, os aglomerados resultantes são caracterizados com base nos padrões pelos quais eles são

formados. Na sequência, pode-se determinar a qual grupo é mais provável pertencer um novo dado com um padrão desconhecido, e caracterizá-lo com base nos padrões do respectivo agrupamento.

Neste trabalho, a Análise de Agrupamentos será utilizada para classificar os tipos de precipitação utilizando dados de radar meteorológico e identificar, nestes dados, a região afetada por banda brilhante.

Três técnicas foram testadas com este intuito: *k-means*, *c-means* e a técnica CLARA (do inglês, *Clustering for Large Applications*), que estende a abordagem *k-medoid* para um grande número de objetos. Porém, o *c-means* obteve resultado semelhante ao *k-means*, já o CLARA obteve resultados muito inferiores. Outros fatores que limitaram o uso de outras técnicas foram a extensa quantidade de dados e o número fixo de grupos desejados. Assim, optou-se por apresentar apenas a técnica *k-means*, e sua descrição é feita a seguir.

3.4 ALGORITMO K-MEANS

O *k-means* é uma técnica de Análise de Agrupamentos do tipo *hard clustering*, que permite dividir um conjunto de dados em um determinado número k de grupos, baseado na dissimilaridade entre os dados. Uma vantagem principal deste algoritmo é a sua simplicidade computacional, o que o torna um bom candidato para uma variedade de aplicações (THEODORIDIS; KOUTROUMBAS, 2009).

A técnica inicia com a escolha do número de grupos e o posicionamento do centro de cada grupo no espaço de atributos. Esse centro é chamado de centroide, e é o ponto médio do grupo. Após a definição da localização destes centroides, cada dado é associado ao centroide mais próximo através do cálculo do quadrado da distância euclidiana.

Quando todos os dados já tiverem sido associados a algum centroide, estes são recalculados (SOUZA, 2008). Este procedimento é repetido inúmeras vezes até que

não haja mais diferença (ou que ela seja mínima) nos centros dos grupos.

Tem-se assim, os seguintes passos para o método:

1. Escolher arbitrariamente k centros iniciais $C = \{C_1, C_2, \dots, C_k\}$;
2. Atribuir cada amostra x_i ($i = 1, \dots, N$) para o centro mais próximo;
3. Determinar os novos centroides de cada grupo;
4. Repetir 2 e 3 até não ocorrer mudanças no centroides de duas iterações consecutivas.

Esta técnica minimiza a seguinte função objetivo:

$$J = \sum_{j=1}^k \sum_{i=1}^N \|x_i^j - C_j\|^2 \quad (21)$$

onde N é o número de realizações da série, k é o número de grupos e $\|x_i^j - C_j\|^2$ é a distância euclidiana entre o dado x_i^j pertencente ao grupo j e o centroide deste mesmo grupo (SOUZA, 2008). Logo, o valor encontrado pela Equação 21 é o somatório das distâncias entre os dados e os centroides dos seus respectivos grupos.

Sua complexidade computacional é da ordem de $O(Nkq)$, onde q é o número de iterações necessárias para a convergência do algoritmo. Pelo fato de k e q serem significativamente menores que N , o *k-means* torna-se elegível para o processamento de grandes conjuntos de dados (THEODORIDIS; KOUTROUMBAS, 2009).

A Figura 14 ilustra a aplicação dos passos para o agrupamento dos dados em dois grupos. O conjunto de dados está representado em (a) pelos pontos em verde, e os dois centroides iniciais, escolhidos aleatoriamente, são representados pelos pontos azul e vermelho, marcados em x. Em (b) cada um dos dados é associado ao seu centroide mais próximo, formando assim dois grupos distintos, um em vermelho, e o outro em azul. Os novos centroides dos grupos são recalculados em (c). O processo se repete até que haja convergência do método, quando não há mais mudança nos centroides dos grupos.

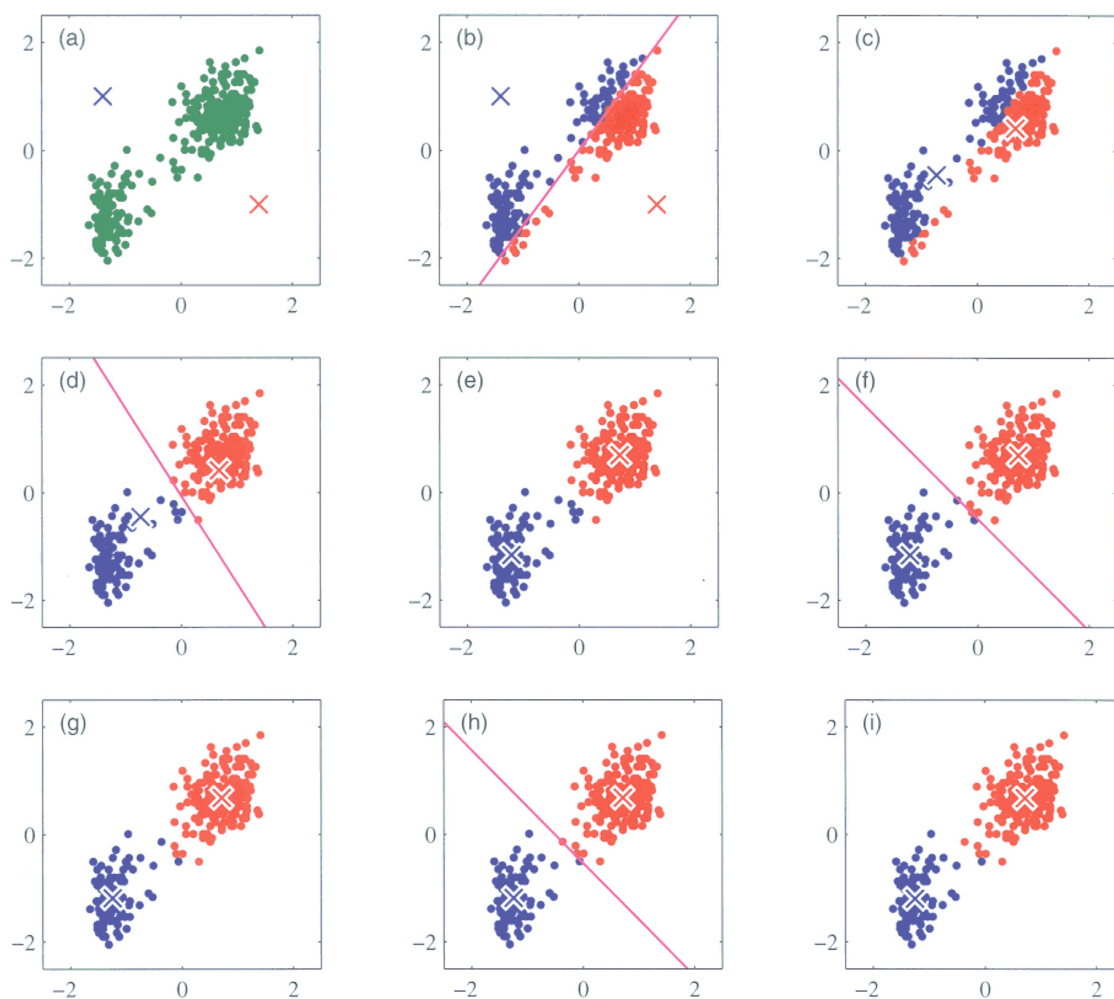


FIGURA 14: Ilustração do algoritmo *k-means*.

FONTE: Damian (2011).

A parte mais importante desta técnica é a sua inicialização, isto é, a localização dos centroides iniciais, pois diferentes centroides irão gerar diferentes resultados (SOUZA, 2008). Isto pode ser observado na Figura 15, onde é feito o agrupamento dos pontos A, B, C, D, E, F e G utilizando o método *k-means* com duas inicializações diferentes dos centroides.

Na Figura 15(a) foram escolhidos como centroides iniciais os pontos A, B e C e os grupos resultantes foram {A}, {B, C} e {D, E, F, G}. Já na Figura 15(b) os centroides iniciais foram A, D e F e obteve-se os grupos {A, B, C}, {D, E} e {F, G}. O valor encontrado pela função objetivo é bem maior para o agrupamento obtido na Figura

15(a) do que para o agrupamento obtido na Figura 15(b).

Sendo assim, Arthur e Vassilvitskii (2006) propuseram uma maneira de escolher centros iniciais para o algoritmo *k-means*. Esse algoritmo chama-se *k-means++*, e é descrito a seguir.

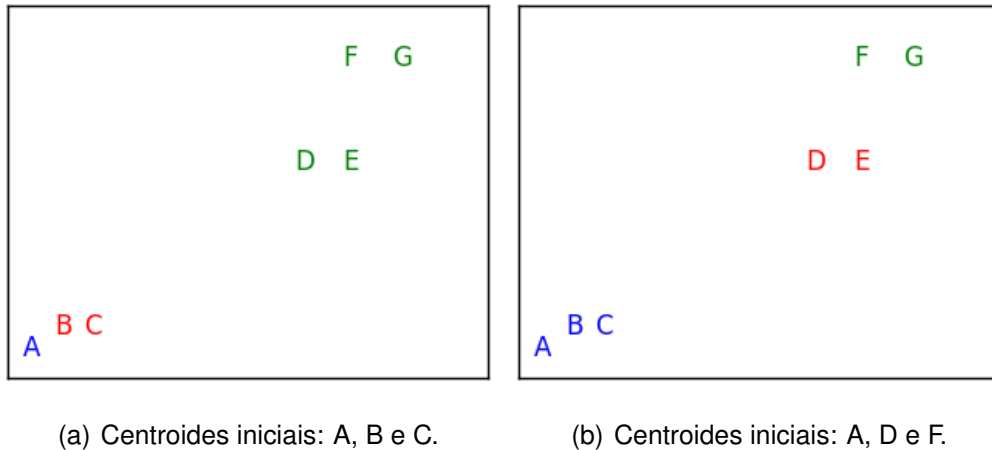


FIGURA 15: Exemplo do algoritmo *k-means* com diferentes inicializações dos centroides.

FONTE: A autora (2014).

3.5 ALGORITMO *K-MEANS++*

O algoritmo *k-means++* escolhe os centroides iniciais baseado pelos seguintes passos:

- 1a. Escolha aleatoriamente um centro inicial C_1 dentro do conjunto de dados;
- 1b. Escolha um novo centro C_i , tal que $C_i = x' \in X$ com probabilidade $\frac{D(x')^2}{\sum_{x \in X} D(x)^2}$;
- 1c. Repita o passo 1b até se obter k centros;
- 2-4. Prossiga com o algoritmo *k-means* padrão.

No passo 1b, $D(x)$ é a distância dos dados para o centroide mais próximo do grupo ao qual está incluso. Este passo atribui uma probabilidade de cada amostra ser um bom centroide inicial.

Segundo Arthur e Vassilvitskii (2006) o novo método produz um algoritmo com melhor desempenho e de forma consistente encontra um agrupamento melhor que o *k-means* padrão.

4 METODOLOGIA

4.1 DADOS

A correção da banda brilhante será aplicada em dados brutos de refletividade, em coordenadas esféricas. Estes dados são obtidos pelo radar meteorológico DWSR-95S, operado pelo SIMEPAR.

Foram selecionados 68 eventos de chuva no Paraná com possível contaminação por banda brilhante, dentro do período de Dezembro/2008 a Setembro/2012. Neste período, a repetição das tarefas de coleta dos dados foi de 10 minutos, com 14 elevações da antena e raio de alcance de 200 km. Os *bins* são coletados a cada 250 m, resultando em uma matriz de dados de 800 x 360 (800 *bins* e 360 feixes) para cada elevação.

Para avaliação dos resultados, serão utilizados dados de precipitação obtidos pelos pluviômetros das estações meteorológicas do SIMEPAR e do INMET (Instituto Nacional de Meteorologia). Estão disponíveis um total de 19 estações do SIMEPAR e 16 estações do INMET no raio de abrangência do radar. A localização de cada estação é mostrada na Figura 16.

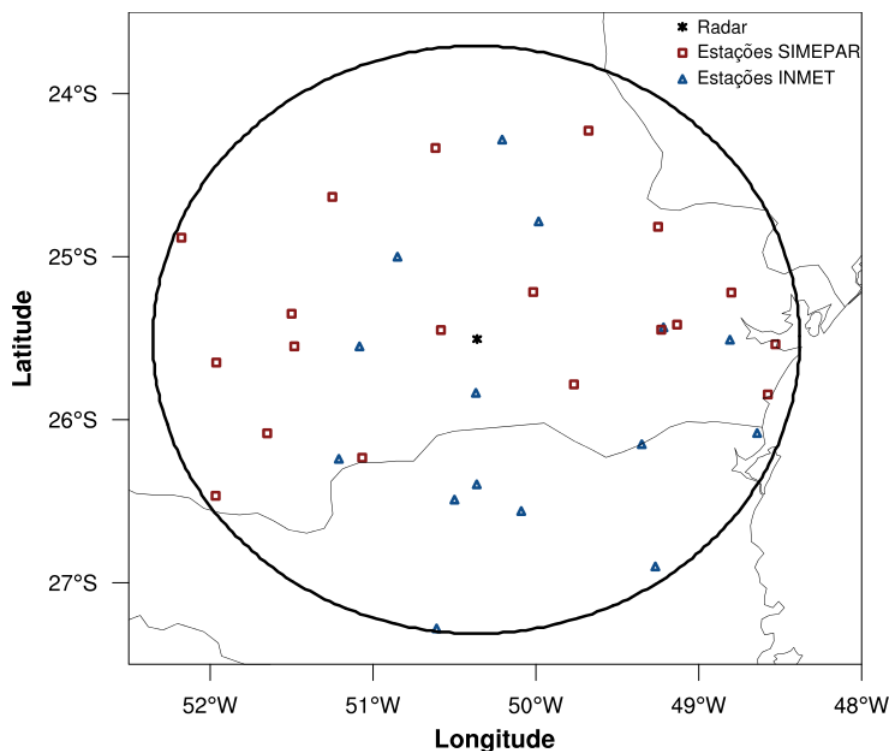


FIGURA 16: Localização das estações meteorológicas do SIMEPAR e INMET.

FONTE: A autora (2014).

4.2 CONTROLE DE QUALIDADE DOS DADOS DE REFLETIVIDADE

Os dados de refletividade podem conter alguns ruídos que não representam precipitação, por isso, há a necessidade de realizar um controle de qualidade.

O radar contém algoritmos que removem ruídos resultantes do processamento do sinal, ecos de terreno, dados espúrios, entre outros processos. Porém, mesmo com esse processamento digital, é necessário um controle de remoção de eco de terreno que possa continuar contaminando os dados.

Neste caso, é necessário retirar os *bins* que apresentem maior contaminação de eco de terreno presentes nas elevações mais baixas (0,5°, 1,0°, 1,5° e 2,0°). A Figura 17 apresenta os PPIs destas elevações com o mapa de eco de terreno utilizado. Este mapa foi obtido após análise de frequência de observações no período de 2000 a 2010, o qual é possível identificar a região predominante contaminada por ecos de terreno.

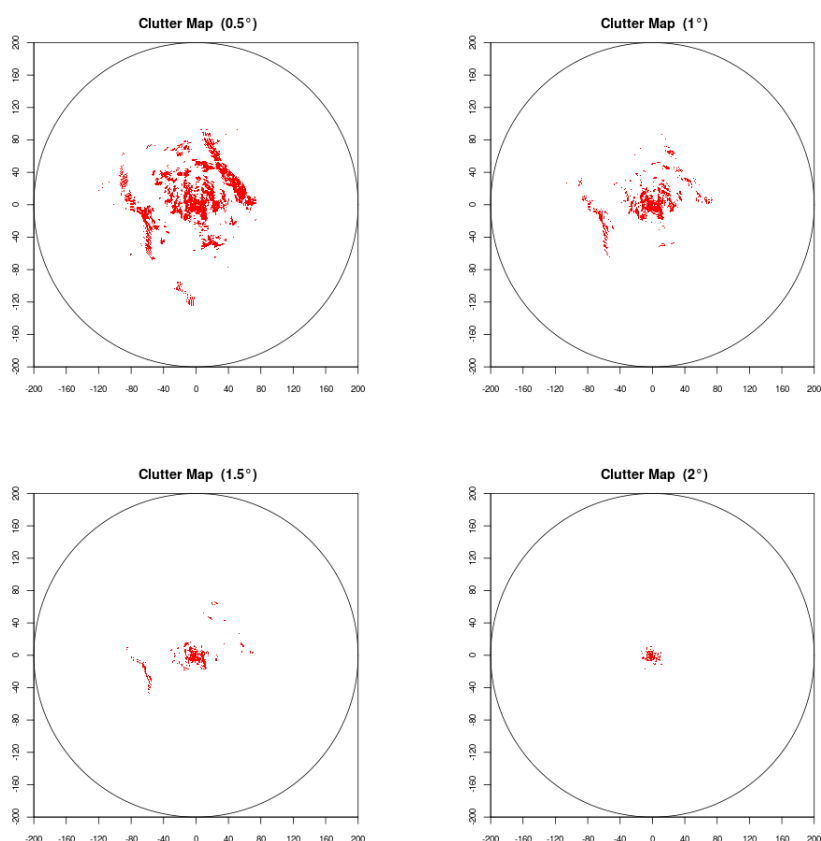


FIGURA 17: Eco de terreno nos dados volumétricos de radar para as elevações de 0,5°, 1,0°, 1,5° e 2,0°.

FONTE: Beneti (2012).

Decidiu-se utilizar também um limiar para a mínima refletividade, neste caso 15 dBZ, uma vez que valores abaixo desta medida representam de garoas à chuvas fracas.

4.3 ISOTERMA DE ZERO GRAU

A isoterma de zero grau, também chamada de nível de congelamento (*freezing level*), é o nível mais baixo em uma sondagem em que é relatada uma temperatura de 0°C. Essa temperatura varia de acordo com o tempo e localização no globo.

Neste trabalho, os valores da isoterma de zero grau foram obtidos a partir dos dados da estação de radiossondagem de Curitiba (código 83840 da OMM - Organi-

zação Meteorológica Mundial). Como a distribuição de estações de radiossondagem no Brasil é muito esparsa, optou-se por selecionar esta única estação para avaliar o perfil de temperatura na área de cobertura do radar (200 km de raio). Foram utilizados os dados das 00 UTC e 12 UTC disponíveis na internet, na página do Departamento de Ciências Atmosféricas da Universidade de Wyoming, nos Estados Unidos (<http://weather.uwyo.edu/upperair/sounding.html>).

Como pode haver diversas falhas na coleta de dados, optou-se por utilizar um perfil vertical de temperatura média sempre que os dados não estivessem disponíveis. Beneti (2012) realizou um estudo no período de Janeiro/2000 a Dezembro/2010 para obter este perfil, apresentado na Figura 18, onde pode-se observar a evolução anual desse perfil no ambiente ao redor do radar e as alturas das isotermas de 0°C, -10°C, -20°C, -30°C e -40°C ao longo do ano.

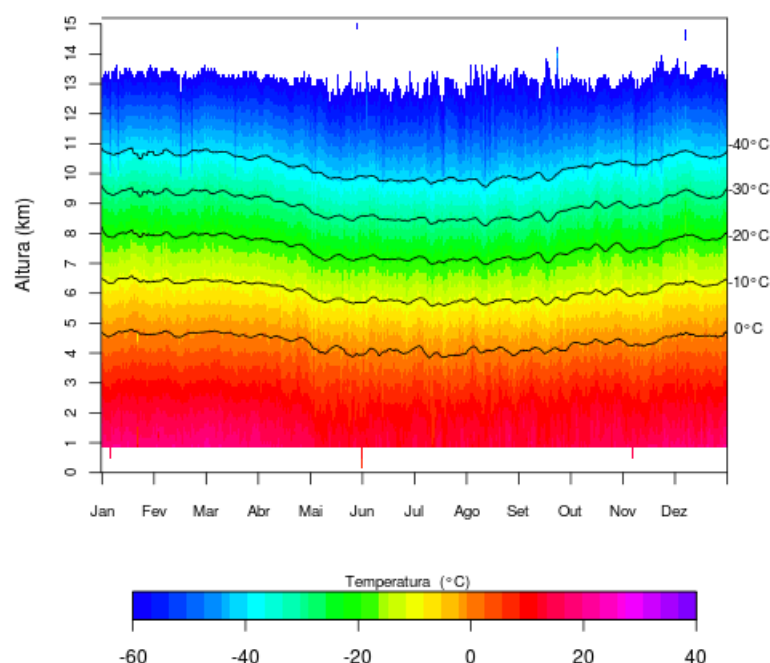


FIGURA 18: Distribuição anual do perfil vertical de temperatura média.

FONTE: Beneti (2012).

4.4 PYTHON

Os métodos utilizados neste trabalho para a identificação da banda brilhante foram desenvolvidos na linguagem de programação *Python*.

O *Python* é uma linguagem de altíssimo nível, orientada a objeto, de tipagem dinâmica e forte, interpretada e interativa (BORGES, 2010). Possui uma sintaxe clara e concisa, que favorece a legibilidade do código fonte, tornando a linguagem uma das mais populares para computação científica.

O *Python* possui uma vasta biblioteca de módulos e pacotes que é distribuída com o interpretador. Existem também disponíveis muitas bibliotecas escritas por terceiros, que implementam diversos recursos além da biblioteca padrão (BORGES, 2010). Uma delas é o *Scikit-learn*, que oferece uma ferramenta simples e eficiente para a mineração e análise de dados. Esta biblioteca possui o módulo *sklearn.cluster.KMeans()*, que agrupa os dados utilizando o método *k-means*. Esta técnica será utilizada para identificar regiões contaminadas pela BB.

Este módulo tem como principais parâmetros de entrada a matriz de dados, a quantidade de grupos desejados e a forma de inicialização dos centroides. Esta inicialização pode ser feita de forma aleatória, através do algoritmo *k-means++* ou pode ser passado um vetor com as posições desejadas dos centroides. Mais informações da biblioteca podem ser encontradas em Pedregosa et al. (2011).

4.5 IDENTIFICAÇÃO DA REGIÃO DE BANDA BRILHANTE

Para a identificação da região de banda brilhante foram utilizados dois métodos distintos, com o objetivo de comparação. Estes métodos são descritos a seguir.

4.5.1 IDENTIFICAÇÃO PELA ESPESSURA APARENTE

Neste método a BB é delimitada através do cálculo de sua espessura aparente. Como a BB está associada à precipitação estratiforme, deve-se primeiramente identificar os dados com essa classificação de precipitação.

Como os modos básicos de precipitação são convectiva e estratiforme, faz-se a classificação dos dados nesses dois tipos de precipitação. Para isso, são realizados dois testes:

- **Teste 1:** marca-se as colunas com valor de $VIL \geq 1 \text{ kg/m}^2$ e $Z_{max} \geq 30 \text{ dBZ}$. Essas colunas são classificadas como convectivas.
- **Teste 2:** em seguida, é feito um agrupamento por dilatação, marcando os pontos com até 4 km das colunas convectivas também como precipitação convectiva.

Então, as colunas não marcadas como convectivas são consideradas estratiformes. Os limites de 1 kg/m^2 e 30 dBZ para o VIL e Z_{max} , respectivamente, foram propostos por Beneti (2012), que desenvolveu seu trabalho na mesma região de interesse deste.

A região estratiforme é então dividida em duas regiões: uma afetada pela BB e outra não afetada. Esta divisão é feita utilizando um primeiro cálculo da espessura da BB:

$$\begin{aligned} BB_{topo} \text{ aparente} &= \text{isoterma de zero grau} + D_1 \\ BB_{base} \text{ aparente} &= 5 \text{ km abaixo do } BB_{topo} \text{ aparente} \end{aligned}$$

O parâmetro D_1 representa a metade da abertura do feixe no ponto onde a isoterma de zero grau intercepta o feixe de menor elevação. Portanto, quanto maior for a altitude da isoterma de zero grau, mais distante do radar está o ponto de intersecção, e maior é o valor de D_1 . Este parâmetro está ilustrado na Figura 19.

O baixo limite inferior (BB_{base}) foi utilizado para abranger as maiores regiões possíveis que possam estar impactadas pela presença de BB, incluindo aquelas associadas

com o derretimento de gelo dentro da precipitação estratiforme atrás de linhas de instabilidade, e para variações da camada de fusão no espaço e tempo.

Após o cálculo da espessura aparente, os dados dentro desta faixa, cujas colunas possuem valores de $Z_{max} \geq 30$ dBZ, são considerados afetados pela banda brilhante.

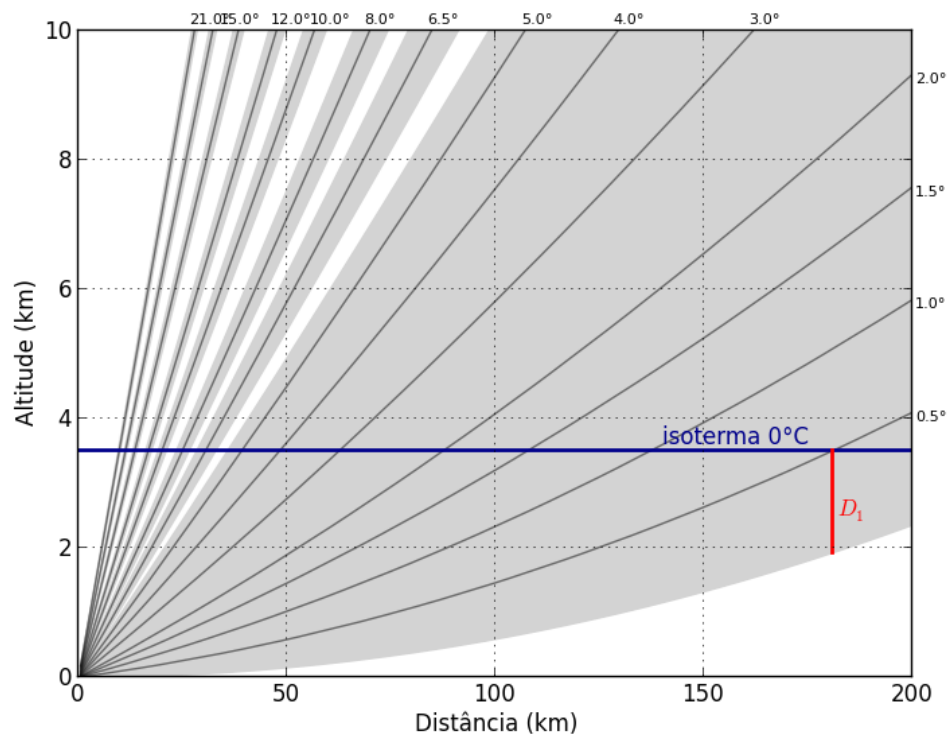


FIGURA 19: Parâmetro D_1 .

FONTE: A autora (2014).

4.5.2 IDENTIFICAÇÃO POR *K-MEANS*

A segunda técnica para a identificação da região de BB utiliza a metodologia *k-means*, que agrupa os dados conforme a quantidade de grupos desejados.

Devido as grandes diferenças entre a estrutura da precipitação convectiva e estratiforme, optou-se por separar os dados em três grupos distintos: uma para a região convectiva, um para a estratiforme e outro para a região de BB. Para a identificação destas três regiões utilizou-se as variáveis:

- Z :

Valor da variável refletividade para o PPI de elevação $0,5^\circ$ (em dBZ). Quanto maior a refletividade, maior a quantidade de gotas de chuva, ou maior o diâmetro das gotas, e, portanto, mais intensa é a precipitação.

- Z_{max} e $h_{Z_{max}}$:

A variável Z_{max} , quando acompanhada da variável $h_{Z_{max}}$, pode indicar uma boa classificação, pois valores de Z_{max} próximos à altura da isoterma de zero grau estão relacionados à existência de banda brilhante. Já quando Z_{max} é encontrada em alturas mais elevadas, significa que a chuva é convectiva.

- h_{30} :

Altura correspondente ao valor de refletividade de 30 dBZ. Alturas elevadas mostram que a tempestade em grandes altitudes é mais intensa, o que está associado à existência de chuva convectiva.

- h_{45} :

Altura correspondente ao valor de refletividade de 45 dBZ. Assim como o h_{30} , ajuda a identificar as regiões de precipitação mais intensa.

- VIL:

Como esta variável representa a integração vertical de toda a água líquida, valores maiores estão relacionados a maior quantidade de precipitação, ou seja, precipitação convectiva.

Essas variáveis formam as colunas da matriz de dados (X) que será um dos parâmetros de entrada para a função `sklearn.cluster.KMeans()`. Esta matriz possui 288000 linhas ($800 \text{ bins} \times 360 \text{ feixes}$) e 6 colunas.

Como elas possuem grandezas e dimensões diferentes, é conveniente normalizá-las. Isso é feito atribuindo uma escala no intervalo $[-1, 1]$ para cada uma das variáveis. Dado um vetor de variáveis, a sua normalização é feita dividindo-se todos os valores

pela norma desse vetor. O objetivo da normalização é evitar que variáveis com escalas significativamente grandes dominem as variáveis de menor escala (DAMIAN, 2011).

Assim, formada a matriz X , aplica-se o método *k-means* para o volume de dados. Como já citado, os outros parâmetros de entrada para a função são o modo de inicialização dos grupos (*k-means++*) e a quantidade de grupos desejados, neste caso, $k = 3$.

4.6 CORREÇÃO DA BANDA BRILHANTE

O método para a correção da banda brilhante utilizado neste trabalho é baseado no método apresentado por Zhang e Qi (2010). Os autores utilizaram o PVR aparente para obter um modelo linear com cinco parâmetros, utilizados para calcular um fator de correção para os dados afetados pela banda brilhante. As vantagens da utilização deste tipo de perfil foram apresentadas na seção 2.5.

4.6.1 CÁLCULO DO PVR

Para cada elevação, um PVR é calculado a partir da média azimutal para cada *range*. Faz-se o cálculo apenas a partir dos dados da área de BB ao invés de todos os dados do radar. Muitas vezes a banda brilhante não está uniformemente distribuída em todas as regiões de cobertura do radar. Assim, se um único PVR médio for calculado a partir de todo o volume, a intensidade do pico da banda brilhante no PVR pode ser reduzida devido à contribuição de dados não afetados. Se este PVR for aplicado para correções em todos os lugares, então haveria subcorreções em áreas de BB e supercorreções em áreas não afetadas. Então, isolar a área de BB do resto da precipitação torna a correção mais eficaz.

Uma vez que o estudo está focado na precipitação estratiforme, presume-se que a estrutura vertical da precipitação é horizontalmente uniforme. Mais especificamente, admite-se que a variação da refletividade ao longo de um feixe é dominado pelo gradiente vertical, em especial nas regiões afetadas pela banda brilhante.

Um procedimento de controle de qualidade é aplicado para excluir valores discrepantes e assegurar um PVR representativo e estável. Uma média móvel de 15 pontos é aplicada no PVR, e todos os pontos que desviam mais do que o dobro do desvio padrão são sinalizadas como ruins. Se mais de 40% dos pontos no PVR são marcados como ruins, então este PVR é considerado inválido. Nessas situações poderia ser utilizado um PVR climatológico. No entanto, PVRs inválidos são geralmente associados com a precipitação dispersa e não homogênea. Portanto, a correção só é aplicada se um PVR válido for obtido.

Após a obtenção de um PVR válido, ele é então parametrizado e cinco parâmetros podem então ser obtidos a partir dele:

- BB_{top} :
Maior altitude encontrada no PVR;
- BB_{pico} :
Altitude correspondente ao maior valor de refletividade encontrado no PVR;
- BB_{base} :
Altitude correspondente ao menor valor de refletividade do PVR, abaixo do BB_{pico} .
Se o menor valor de refletividade for menor que o limite de 29 dBZ, então BB_{base} receberá o valor da altitude correspondente a esse limite. Esse limite é usado para evitar correções excessivas (ZHANG; QI, 2010);
- α :
É obtido pelo método de mínimos quadrados através dos pontos do PVR entre BB_{topo} e BB_{pico} ;
- β :
É obtido também pelo método de mínimos quadrados, porém entre os pontos do PVR entre BB_{pico} e BB_{base} .

A Figura 20 mostra um PVR com esses cinco parâmetros.

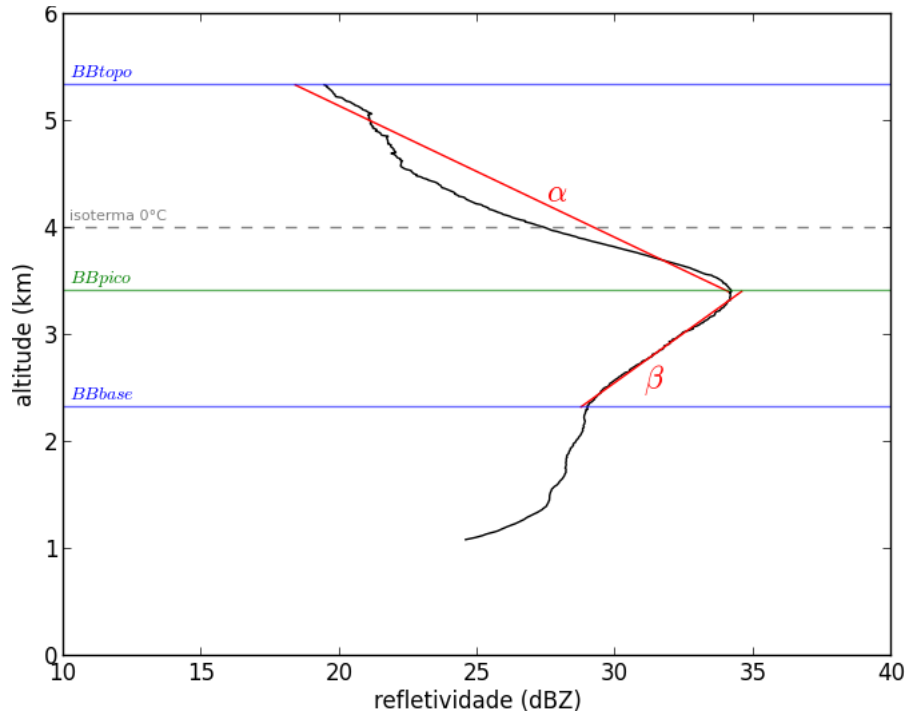


FIGURA 20: Parâmetros obtidos a partir do PVR.

FONTE: A autora (2014).

Nota-se que, em uma estrutura típica de banda brilhante, α é sempre negativo e β é sempre positivo. Caso contrário, a precipitação é considerada não afetada por banda brilhante e nenhuma correção é aplicada.

4.6.2 APLICAÇÃO DA CORREÇÃO

Os parâmetros obtidos a partir do PVR calculado na seção anterior são utilizados para se obter um fator de correção para a região de banda brilhante. Este fator é calculado da seguinte maneira:

$$dbz_a(h) = \begin{cases} \alpha[h(r) - BB_{pico}] + \beta[BB_{pico} - BB_{base}], & \text{se } h(r) > BB_{pico} \\ \beta[h(r) - BB_{base}], & \text{se } h(r) \leq BB_{pico} \end{cases} \quad (22)$$

Assim, o valor de refletividade corrigido (dBZ_c) é obtido de acordo com a equação 18:

$$dBZ_c(\theta, 0) = dBZ_o(\theta, h) - dBZ_a(h), \quad (23)$$

onde dBZ_o é o valor de refletividade original do radar na região da banda brilhante.

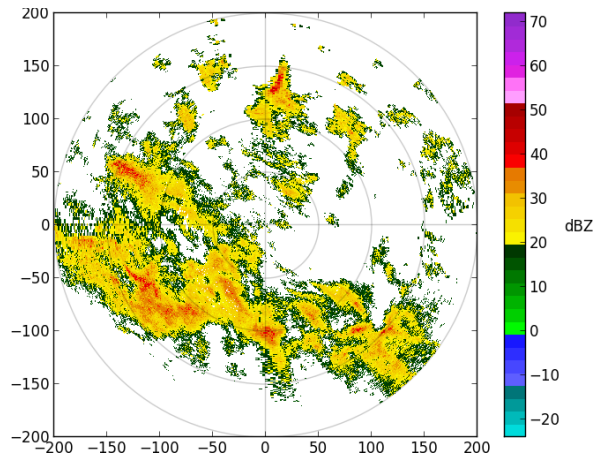
Os dois métodos de identificação da banda brilhante foram aplicados para os casos selecionados, e então, a correção foi utilizada nas regiões indicadas pelos métodos como afetadas. Os resultados obtidos com a correção, bem como uma comparação com valores de precipitação observada são apresentados no próximo capítulo.

5 RESULTADOS

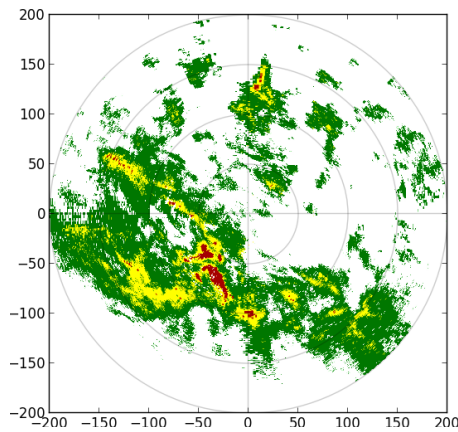
Para cada caso selecionado foram aplicados os métodos de identificação e correção a todos os volumes de dados coletados no intervalo de uma hora de ocorrência do evento. A correção é aplicada para todas as elevações dos volumes de dados, porém, para comparação com dados pluviométricos utilizou-se apenas a primeira elevação ($0,5^\circ$), por ser a mais próxima da superfície terrestre, e por consequência, mais próxima dos pluviômetros.

A seguir, são apresentados os resultados obtidos para os eventos ocorridos em 16/07/2010 às 3:00 (Figura 21), 22/11/2010 às 5:00 (Figura 22) e 12/04/2011 às 18:00 (Figura 23). Os PPIs mostrados são da elevação de $0,5^\circ$ para o primeiro volume de varredura da hora. Em (a) é mostrado o PPI sem correção, em (b) e (d) são mostradas as classificações dos hidrometeoros pelo método de identificação pelo cálculo da espessura aparente e pelo método por *k-means*, respectivamente. Em (c) e (e) tem-se os PPIs dos dados de refletividade após a correção.

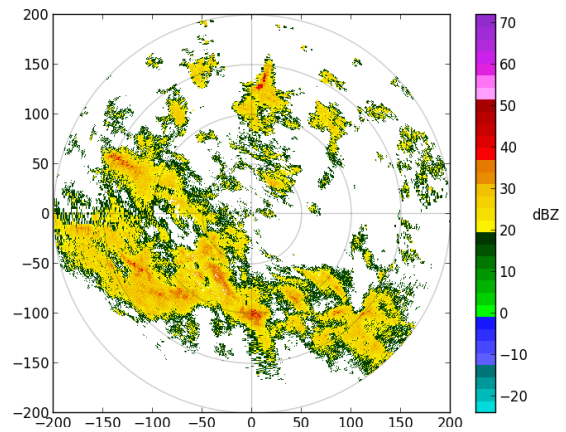
A região de gelo foi identificada posteriormente à utilização dos métodos de identificação da banda brilhante. É considerado como gelo os *bins* que estão completamente acima da isoterma de zero grau, ou seja, os *bins* que possuem o fundo o feixe acima desta isoterma.



(a) PPI sem correção.



(b) Espessura aparente.



(c) Valores corrigidos (espessura aparente).

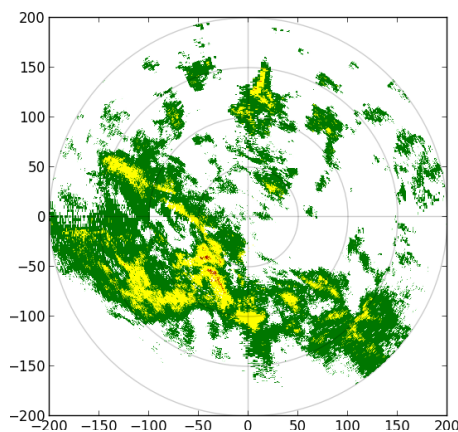
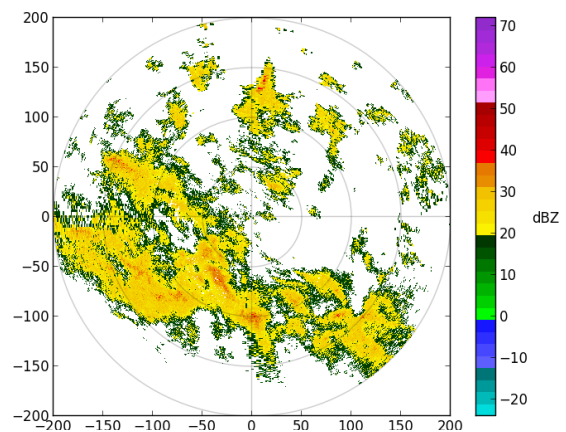
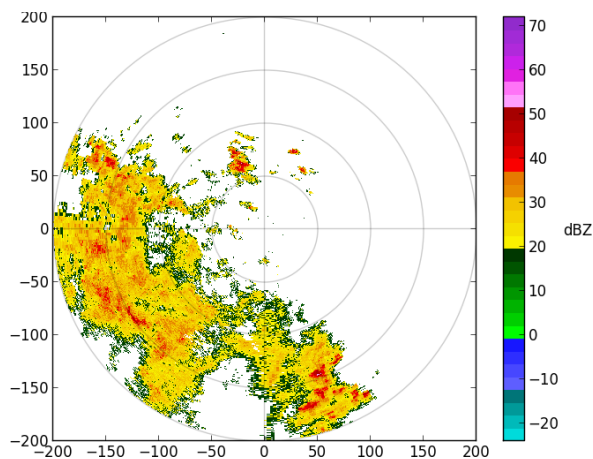
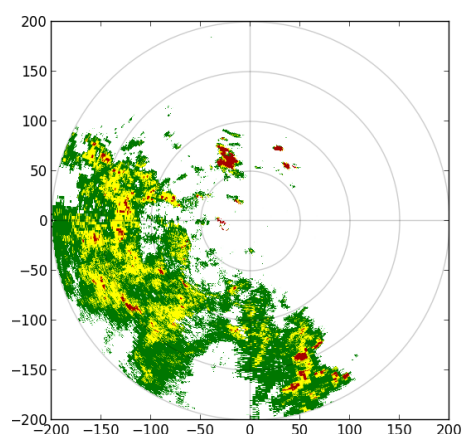
(d) *k-means*.(e) Valores corrigidos (*k-means*).

FIGURA 21: PPI da elevação de 0,5° do dia 16/07/2010 às 03:00. Em (a) tem-se o PPI dos dados originais, em (b) e (d) a identificação da banda brilhante pelo espessura aparente e *k-means*, respectivamente, e em (c) e (e) o PPI após a correção para cada método.

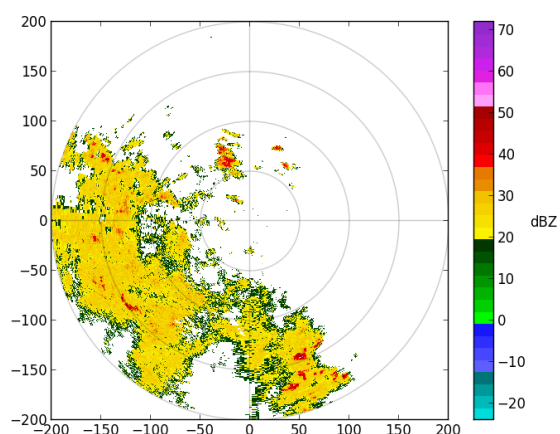
FONTE: A autora (2014).



(a) PPI sem correção.



(b) Espessura aparente.



(c) Valores corrigidos (espessura aparente).

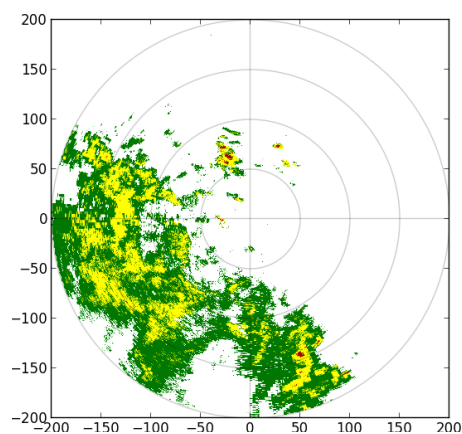
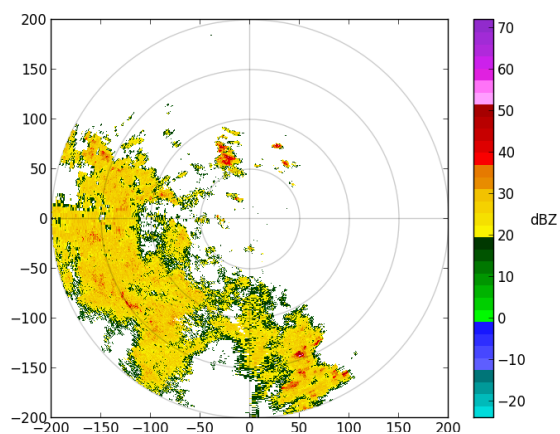
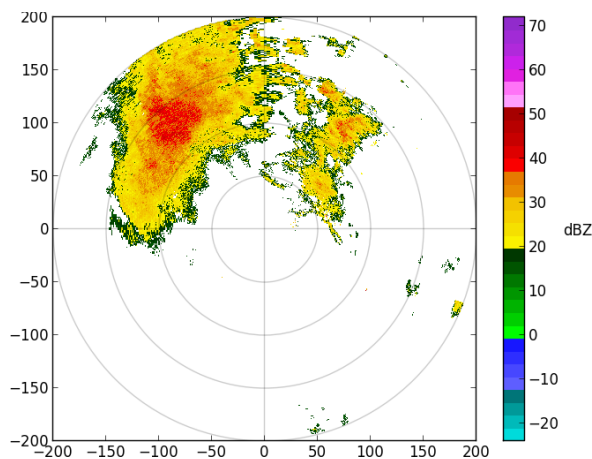
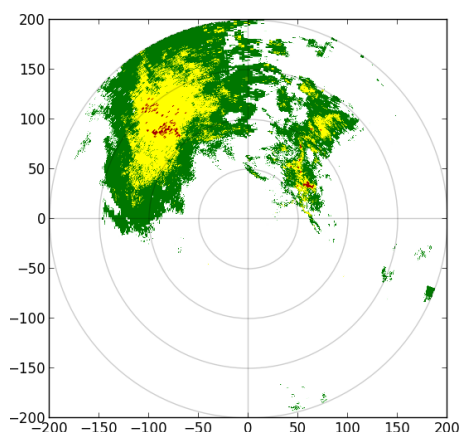
(d) *k-means*.(e) Valores corrigidos (*k-means*).

FIGURA 22: PPI da elevação de $0,5^\circ$ do dia 22/11/2010 às 5:00. Em (a) tem-se o PPI dos dados originais, em (b) e (d) a identificação da banda brilhante pelo espessura aparente e *k-means*, respectivamente, e em (c) e (e) o PPI após a correção para cada método.

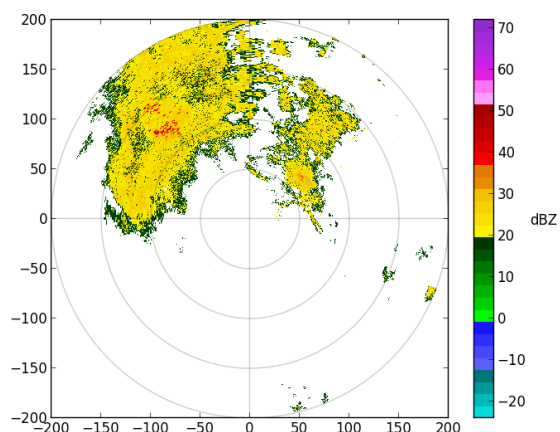
FONTE: A autora (2014).



(a) PPI sem correção.



(b) Espessura aparente.



(c) Valores corrigidos (espessura aparente).

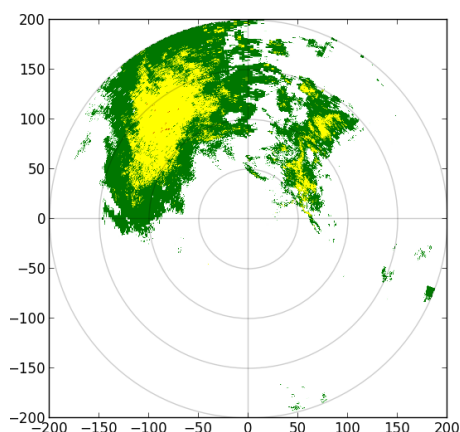
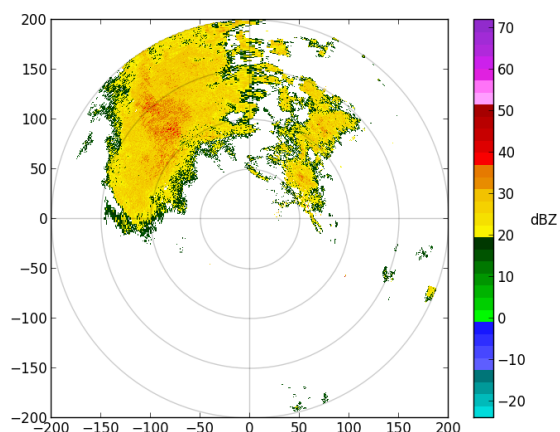
(d) *k-means*.(e) Valores corrigidos (*k-means*).

FIGURA 23: PPI da elevação de 0,5° do dia 12/04/2011 às 18:00. Em (a) tem-se o PPI dos dados originais, em (b) e (d) a identificação da banda brilhante pelo espessura aparente e *k-means*, respectivamente, e em (c) e (e) o PPI após a correção para cada método.

FONTE: A autora (2014).

5.1 AVALIAÇÃO DOS MÉTODOS

Devido a diferença de intervalo de tempo no qual são coletados os dados do radar (10 minutos) e dos pluviômetros (15 minutos), optou-se por utilizar os valores de precipitação no intervalo de uma hora. Para os dados de precipitação obtidos pelos pluviômetros, os valores são acumulados. Já para os dados do radar, decidiu-se por utilizar o valor de refletividade máxima dos PPIs de $0,5^\circ$ em cada hora medida. Este valor de refletividade máxima é então convertida para taxa de precipitação utilizando a relação $Z-R$: $Z = 52 R^{2,7}$, apresentada na seção 2.2.

As comparações são realizadas em pontos específicos onde estão instalados os pluviômetros. Devido ao filtro utilizado, que retira os valores de refletividade menores que 15 dBZ, elimina-se também os dados de pluviômetros dos pontos correspondentes.

Para comparação dos métodos, calculou-se o Viés e a Raiz do Erro Médio Quadrático (REMQ) para os 68 casos.

O Viés é uma variável que indica a direção média dos desvios de um conjunto a partir de um conjunto de valores observados. Quando o valor é positivo indicam superestimativas, enquanto que valores negativos indicam subestimativas. Ele é definido por:

$$\text{Viés} = \frac{1}{N} \sum_{i=1}^N (P_i - P_i^*) \quad (24)$$

onde N é a quantidade de amostras, P é a taxa de precipitação calculada com a relação $Z-R$ e P^* é a precipitação medida pelos pluviômetros.

O REMQ é uma das medidas dos padrões de acurácia das estimativas que indica a magnitude média do erro. Este índice será mais influenciado quando no conjunto de estimativas verificadas existirem erros de maior magnitude, mesmo que sejam poucos, do que quando ocorrem muitos erros pequenos, já que ao calcular a potência dois os

erros maiores serão realçados (SANTOS, 2014). A REMQ é definida por:

$$\text{REMQ} = \sqrt{\frac{1}{N} \sum_{i=1}^N (P_i - P_i^*)^2}. \quad (25)$$

As Figuras 24, 25 e 26 apresentam a taxa de precipitação obtidas com o cálculo da relação $Z-R$ para os três eventos mostrados anteriormente. Observa-se que os PPIs estão acumulados por hora. Nestas Figuras, tem-se também os valores de precipitação observados para cada estação, nos pontos onde se tem pares de pontos válidos. Para cada um destes eventos, é apresentada uma tabela com os valores da taxa de precipitação dos dados originais, dos dados corrigidos pelos dois métodos e da precipitação observada para os pares de pontos apresentados nas imagens e outra tabela com o cálculo do Viés e REMQ.

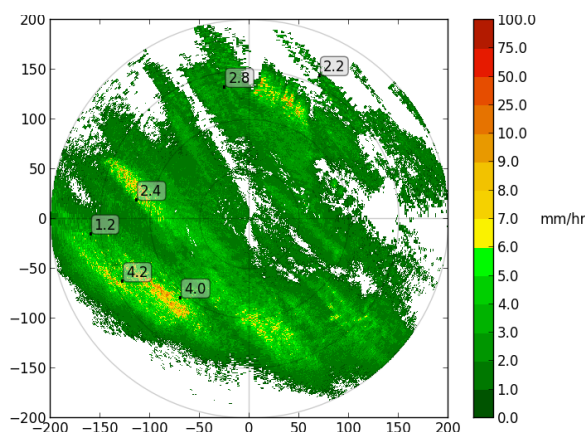
Nota-se que, apesar de se utilizar a mesma correção para ambos os métodos de identificação, os valores da taxa de precipitação após a correção são diferentes, mesmo para os pontos igualmente classificados pelos métodos como afetados. Isso acontece pois, para cada método utilizado tem-se uma região diferente classificada como BB. Isto afeta o valor dos parâmetros obtidos através do PVR desta região. E consequentemente, afeta o cálculo do valor do fator de correção utilizado.

As Figuras 27 e 28 mostram, respectivamente, os valores do Viés e do REMQ para cada caso. Nota-se que em 8 casos houveram correções excessivas que subestimaram o cálculo da taxa de precipitação. Porém, no restante dos casos, obteve-se boa redução dos valores tanto para o Viés, quanto para o REMQ.

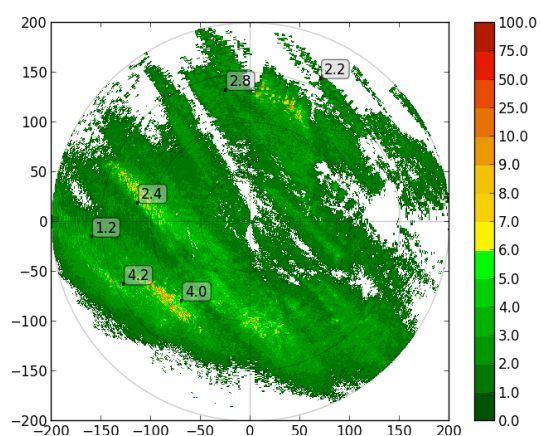
Além dos erros para cada caso, foram calculados os erros para o número total de pontos obtidos dos casos selecionados, mostrados na Tabela 1.

TABELA 1: Viés e REMQ para o total de dados obtidos dos 68 eventos.

	Viés	REMQ
Dados originais	2,0	3,2
Espessura aparente	1,2	2,7
<i>K-means</i>	1,0	2,4



(a) Sem correção.



(b) Espessura aparente.

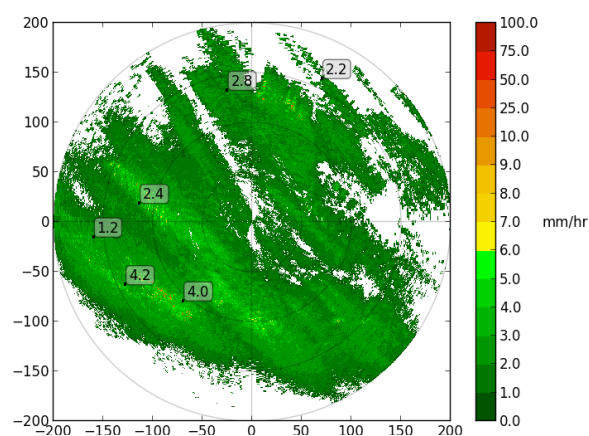
(c) *k-means*.

FIGURA 24: Taxa de precipitação (16/07/2010 às 03:00).

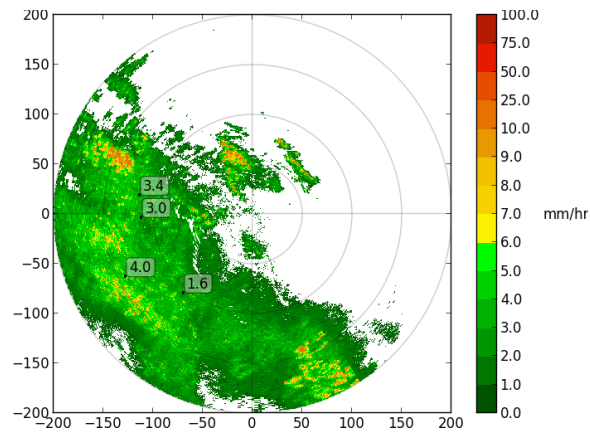
FONTE: A autora (2014).

TABELA 2: Valores da taxa de precipitação (em mm/hr) para 16/07/2010 às 3:00.

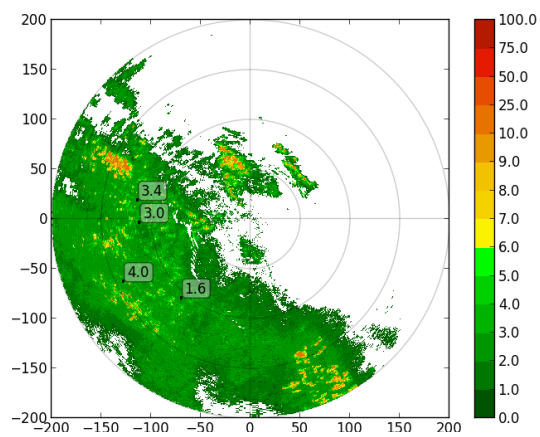
Dados originais	Espessura aparente	<i>K-means</i>	Valor observado
8,0	5,8	4,5	4,2
5,4	3,1	3,1	2,4
2,3	2,3	2,3	2,2
3,4	2,9	3,3	1,2
7,0	6,3	4,2	4,0
2,7	2,7	2,7	2,8

TABELA 3: Valores de Viés e REMQ para 16/07/2010 às 3:00.

	Viés	REMQ
Dados originais	2,0	2,5
Espessura aparente	1,1	1,4
<i>K-means</i>	0,5	0,9



(a) Sem correção.



(b) Espessura aparente.

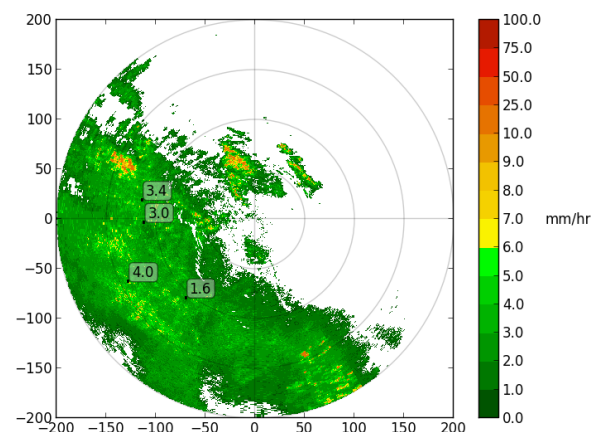
(c) *k-means*.

FIGURA 25: Taxa de precipitação (22/11/2010 às 05:00).

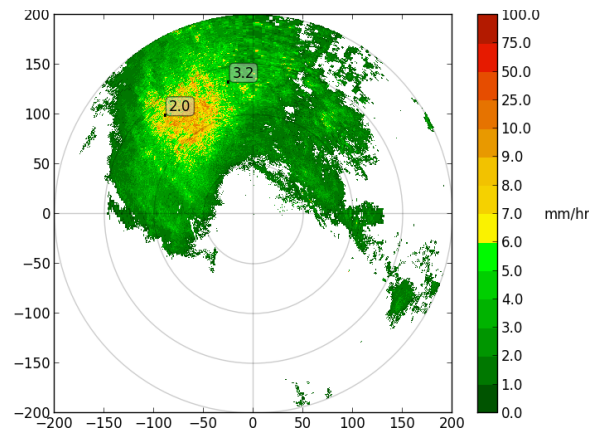
FONTE: A autora (2014).

TABELA 4: Valores da taxa de precipitação (em mm/hr) para 22/11/2010 às 5:00.

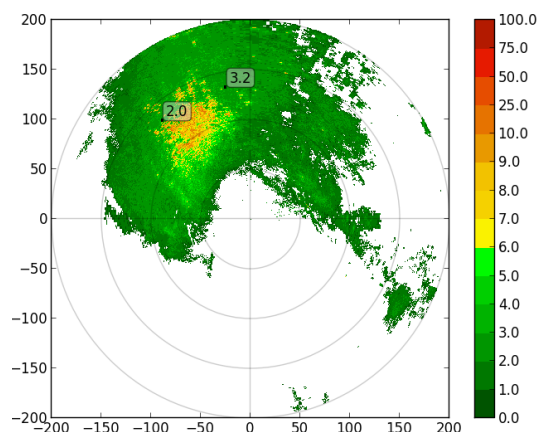
Dados originais	Espessura aparente	<i>K-means</i>	Valor observado
4,2	3,0	3,5	3,0
8,3	4,4	5,9	4,0
4,2	3,4	3,8	3,4
3,3	2,9	2,9	1,6

TABELA 5: Valores de Viés e REMQ para 22/11/2010 às 5:00.

	Viés	REMQ
Dados originais	2,0	2,4
Espessura aparente	0,4	0,7
<i>K-means</i>	1,0	1,2



(a) Sem correção.



(b) Espessura aparente.

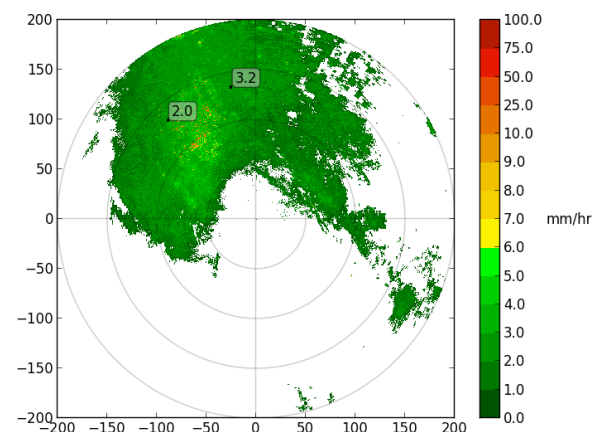
(c) *k-means*.

FIGURA 26: Taxa de precipitação (12/04/2011 às 18:00).

FONTE: A autora (2014).

TABELA 6: Valores da taxa de precipitação (em mm/hr) para 12/04/2011 às 18:00.

Dados originais	Espessura aparente	<i>K-means</i>	Valor observado
8,7	4,5	3,8	2,0
7,6	7,6	3,3	3,2

TABELA 7: Valores de Viés e REMQ para 12/04/2011 às 18:00.

	Viés	REMQ
Dados originais	5,5	5,7
Espessura aparente	3,4	3,6
<i>K-means</i>	0,9	1,3

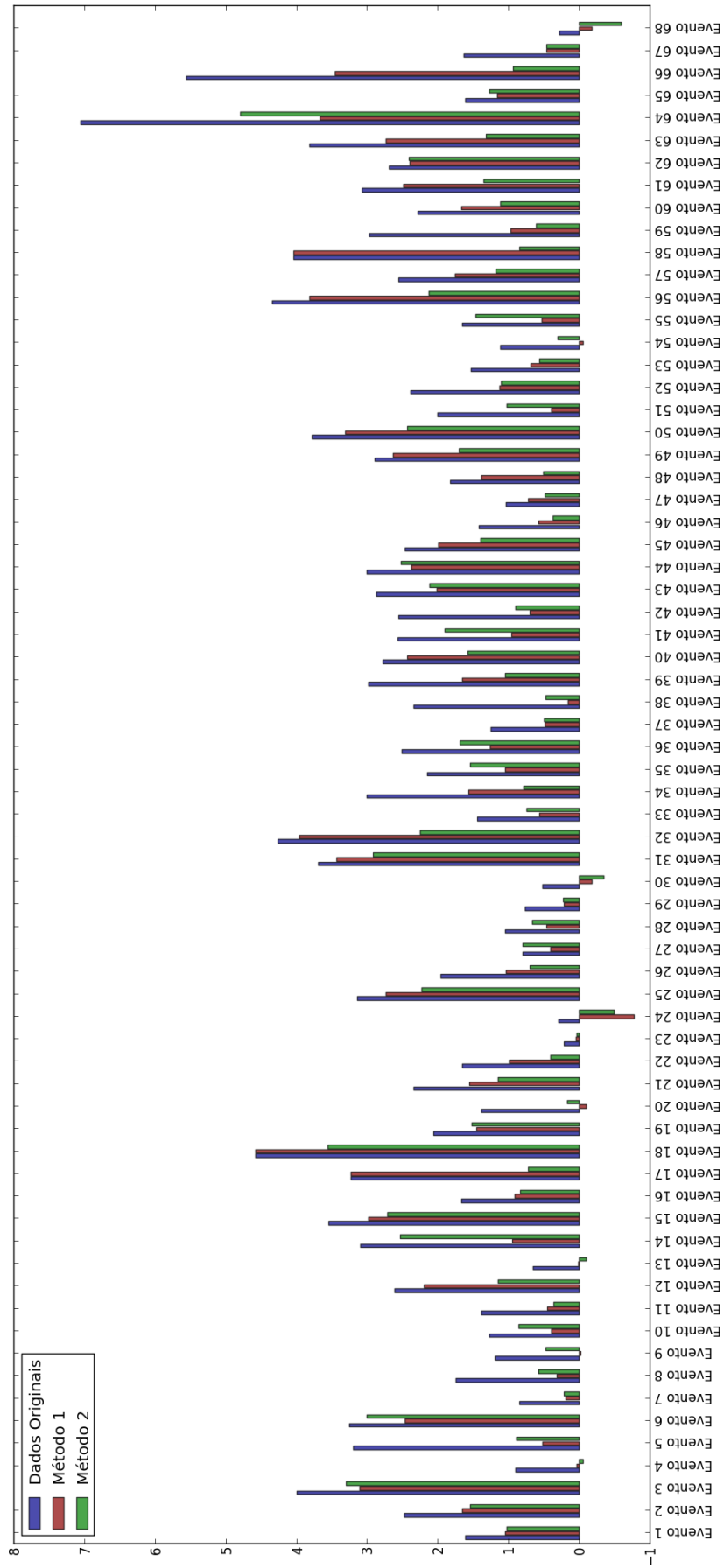


FIGURA 27 : Viés para cada evento.

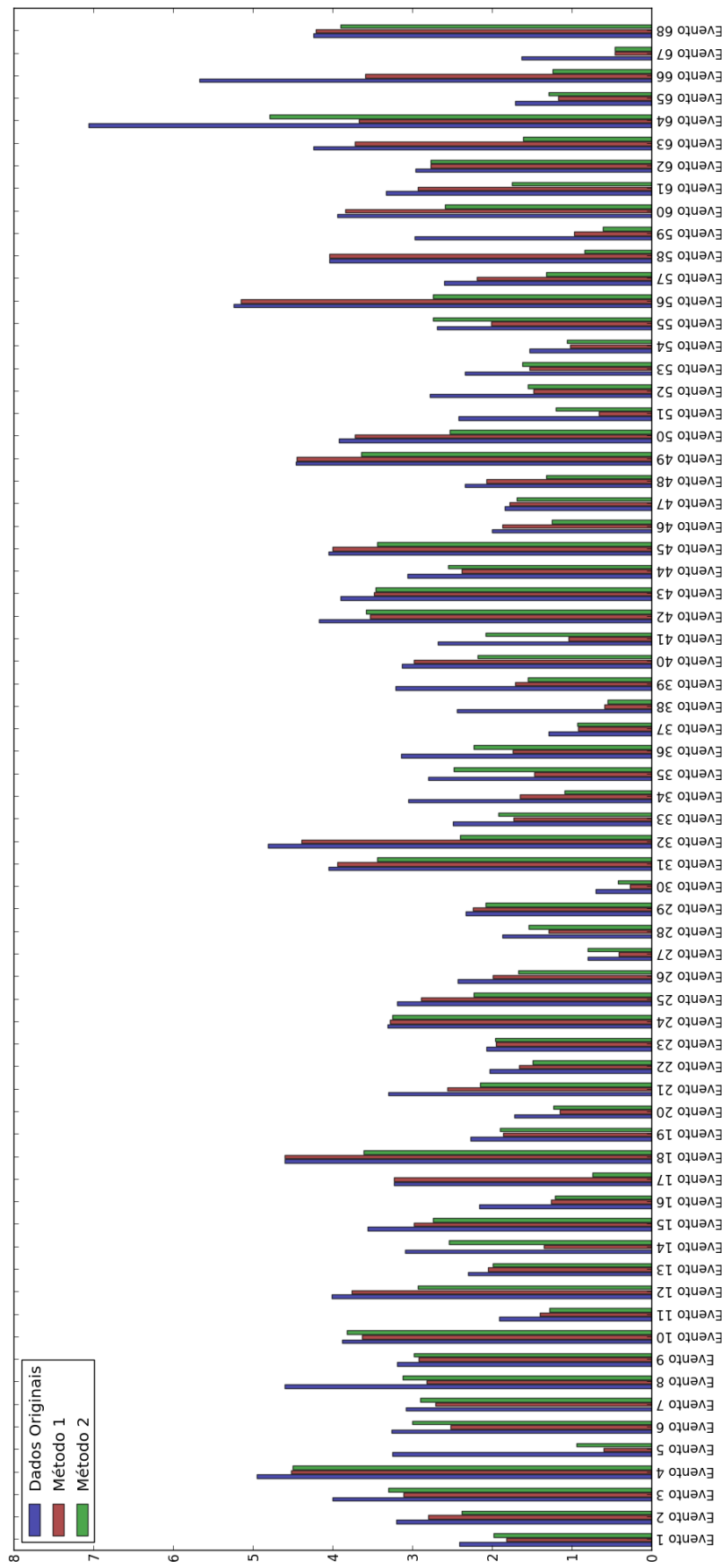


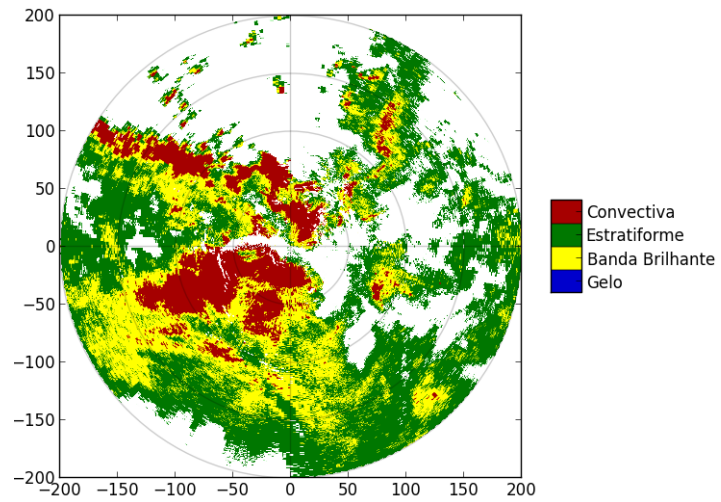
FIGURA 28: REMQ para cada evento.

Na comparação entre os dois métodos utilizados para a identificação, o *k-means* obteve, em geral, os melhores resultados. Nas Figuras 21, 22 e 23, pode-se observar que a principal diferença entre os métodos é a região identificada como convectiva. Os elevados valores da camada de banda brilhante podem influenciar no cálculo do VIL, assim, algumas regiões podem ser erroneamente identificadas como convectivas pelo primeiro método.

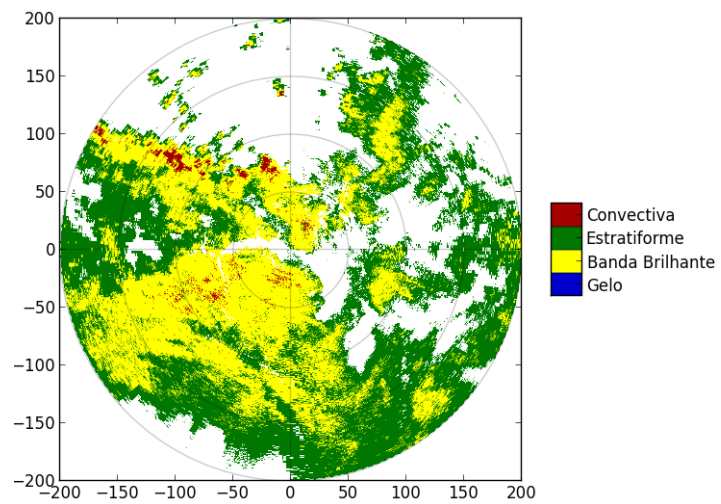
Um exemplo é mostrado na Figura 29, onde tem-se os PPIs com a classificação dos hidrometeoros feita pelos dois métodos, para o dia 15/10/2010 às 16:00 (Evento 46). Nota-se que o método da espessura aparente classifica como convectivo uma região bem mais expressiva que o *k-means*. Selecionou-se então, uma pequena área da região classificada como convectiva no método da espessura aparente (mostrado em cinza na Figura 30(a)) para o cálculo do PVR, mostrado em 30(b). Este PVR é nitidamente um perfil de dados afetados pela BB e possui um pico no valor de refletividade em torno da altitude de 3,5 km, logo abaixo da isoterma de zero grau.

Dois cortes verticais foram realizados nos dados originais, nas linhas indicadas na Figura 30(a), no sentido de *a* para *b* e no sentido de *c* para *d*. No primeiro corte (Figura 30(c)), a região indicada pelo quadrado é característica de uma precipitação convectiva, com valores mais elevados e núcleos de extensão vertical, conforme ilustrado na Figura 12. Isto não é observado na região indicada pela elipse, que foi classificada como convectiva pelo método da espessura aparente. O segundo corte (Figura 30(d)) foi feito na direção radial do radar, onde também não são observadas as características de uma precipitação convectiva. Ou seja, os dados foram erroneamente classificados como convectivos.

No *k-means*, a região verdadeira de precipitação convectiva ficou subestimada. Isto poderia ser corrigido com um agrupamento por dilatação, após a classificação do *k-means*.



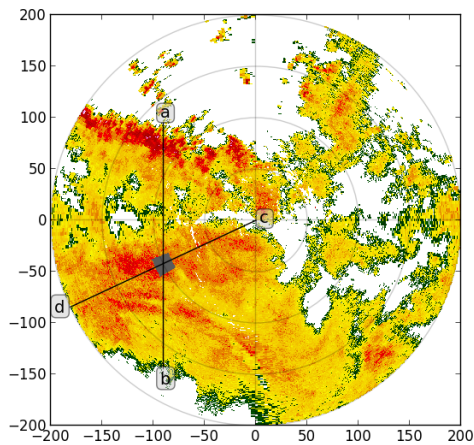
(a) Espessura aparente.



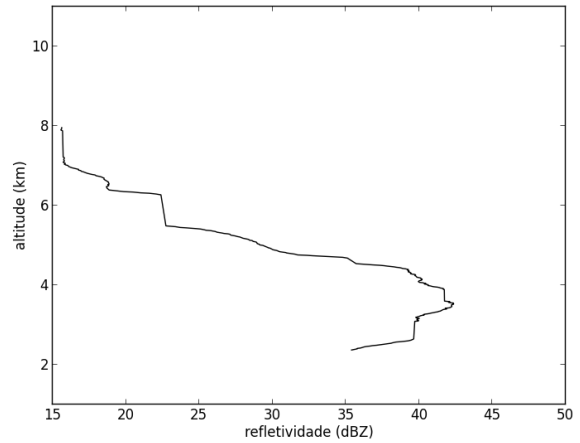
(b) *K-means*.

FIGURA 29: Classificação dos hidrometeoros pelos dois métodos (15/10/2010 às 16:00).

FONTE: A autora (2014).



(a) PPI dos dados originais.



(b) PVR dos dados indicados em cinza.

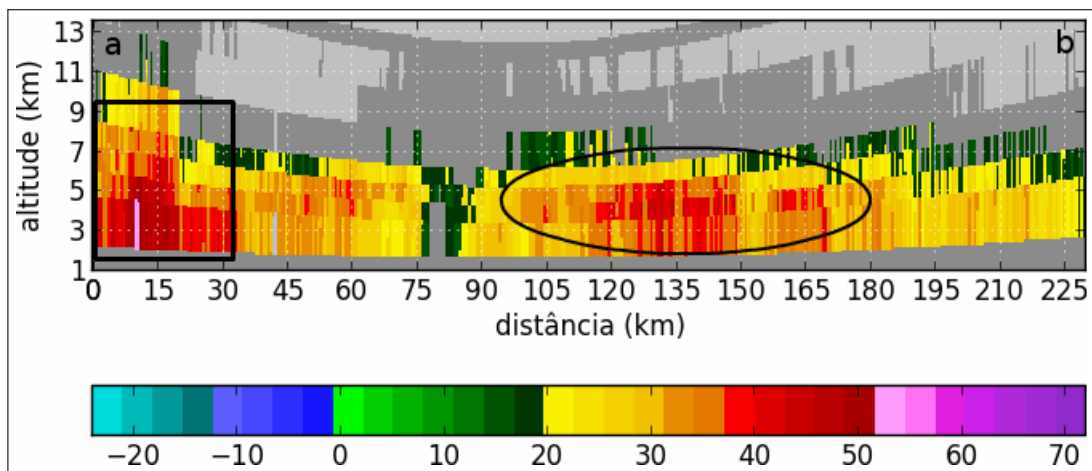
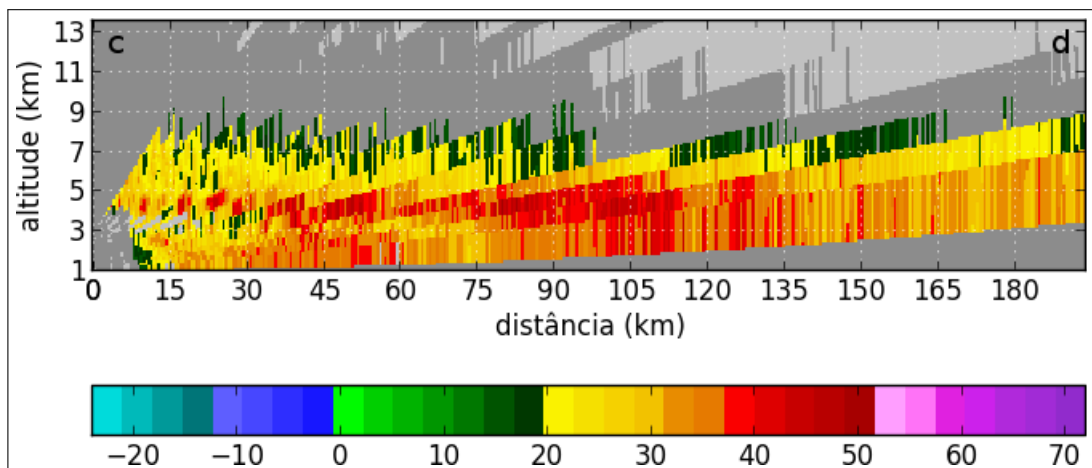
(c) Corte realizado no sentido de *a* para *b*.(d) Corte realizado no sentido de *c* para *d*.

FIGURA 30: PVR e cortes verticais (15/10/2010 às 16:00). Os cortes foram feitos no sentido de *a* para *b*. Em (c) o quadrado indica uma região convectiva e a elipse indica uma região de BB.

FONTE: A autora (2014).

6 CONCLUSÕES

A estimativa da precipitação de chuva com a melhor acurácia possível é fundamental para uma adequada e realística representação da distribuição espacial da precipitação. Uma maneira de estimar a precipitação é a utilização de dados de refletividade obtidos por radares meteorológicos. Assim, há a necessidade de eliminar ou corrigir fatores nestes dados que podem prejudicar esta estimativa de precipitação.

Um destes fatores é a banda brilhante, que pode gerar uma superestimativa no cálculo da taxa de precipitação. O foco deste trabalho era justamente identificar esta região para que uma correção pudesse ser aplicada.

Dois métodos para a identificação da BB foram apresentados. No primeiro, utiliza-se uma classificação da precipitação em convectiva e estratiforme utilizando VIL, para então delimitar a espessura da BB dentro da região estratiforme. Este método utiliza decisões subjetivas para delimitar a BB, assim, utilizou-se também a técnica de agrupamentos *k-means*, que baseia-se na similaridade dos dados para agrupá-los.

Foram selecionados 68 eventos de chuva no raio de alcance do radar, cujos dados estavam contaminados por BB. Os dois métodos foram então aplicados nestes dados e, após a identificação da região afetada, utilizou-se uma correção baseada no perfil vertical da refletividade.

Para avaliar os resultados, converteu-se os dados de refletividade para taxa de precipitação utilizando a relação *Z-R*: $Z = 52 R^{2,7}$. Em seguida, esses valores foram comparados com a precipitação medida por pluviômetros localizados no raio de alcance do radar.

Os dois métodos obtiveram melhores resultados quando comparados com os dados originais. Porém, na comparação entre os dois métodos, o *k-means* mostrou, em

geral, os melhores resultados.

Observou-se que, a maior diferença entre os dois métodos é a identificação da precipitação convectiva. Os elevados valores da camada de banda brilhante podem influenciar no cálculo do VIL, assim, algumas regiões podem ser erroneamente identificadas como convectivas pelo primeiro método.

6.1 SUGESTÕES PARA TRABALHOS FUTUROS

Como sugestões para trabalhos futuros, propõe-se:

- Melhorar a classificação da precipitação em convectiva e estratiforme utilizando outras técnicas, como por exemplo, redes neurais.
- Utilizar os dados do radar meteorológico instalado no município de Cascavel-PR, o qual possui dupla polarização, sendo capaz de identificar a banda brilhante. Seus dados podem ser usados para ajudar na identificação da BB nos dados do radar de Teixeira Soares, uma vez que as áreas de instabilidade podem passar pelo raio de alcance da área quantitativa dos dois radares.
- Aplicar as técnicas para dados de outros radares com polarização simples existentes no Brasil.

REFERÊNCIAS

- ALBRECHT, R. I.; DIAS, M. A. F. da S. Características microfísicas da precipitação convectiva e estratiforme associadas à oscilação de larga-escala no sudoeste da Amazônia. In: **XIII Congresso Brasileiro de Meteorologia**. [S.l.: s.n.], 2004.
- ALMEIDA, M. V. de. **Estimativa de precipitação via radar meteorológico do pico do coto**. Dissertação (Mestrado em Engenharia Oceânica) — Programa de Pós-Graduação de Engenharia, Universidade Federal do Rio de Janeiro, 2002.
- AMBURN, S. A.; WOLF, P. L. VIL density as a hail indicator. **Weather and Forecasting**, v. 12, p. 473–478, 1997.
- ANAGNOSTOU, E. N. A convective/stratiform precipitation classification algorithm for volume scanning weather radar observations. **Meteorological Applications**, v. 11, p. 291–300, 2004.
- ANDRIEU, H.; CREUTIN, J. D. Identification of vertical profiles of radar reflectivity for hydrological applications using an inverse method. Part I: Formulation. **Journal of Applied Meteorology**, v. 34, p. 225–239, 1995.
- ARTHUR, D.; VASSILVITSKII, S. **K-means++: the advantages of careful seeding**. [S.l.], 2006. Disponível em: <<http://ilpubs.stanford.edu:8090/778/>>.
- BENETI, C.; NOZU, I.; SARAIVA, E. A. Monitoramento da precipitação e de eventos de tempo severo com radar meteorológico no Estado do Paraná. In: **XII Cong. Bras. de Meteorologia**. [S.l.: s.n.], 1998.
- BENETI, C. A. A. **Caracterização hidrodinâmica e elétrica de sistemas convectivos de mesoescala**. Tese (Doutorado em Meteorologia) — Departamento de Ciências Atmosféricas, Universidade de São Paulo, 2012.
- BORGES, L. E. **Python para desenvolvedores**. 2. ed. [s.n.], 2010. ISBN 978-85-909451-1-6. Disponível em: <<http://ark4n.wordpress.com/python/>>.
- DAMIAN, E. A. **Dois metodologias aplicadas à classificação de precipitação convectiva e estratiforme com radar meteorológico: SVM e K-means**. Dissertação (Mestrado) — Programa de Pós-Graduação em Métodos Numéricos para Engenharia, Universidade Federal do Paraná, 2011.
- DOVIAK, R. J.; ZRNIC, D. S. **Doppler radar and weather observations**. 2. ed. [S.l.]: Dover Publications, 1993. ISBN 0-486-45060-0.
- FABRY, F.; ZAWADZKI, I. Long-term radar observations of the melting layer of precipitation and their interpretation. **Journal of the Atmospheric Sciences**, v. 52, n. 7, p. 838–851, 1995.

FILHO, K. L. da R.; CONDE, F.; ANDRIOLI, C. P. Correção em tempo real da precipitação estimada por um radar meteorológico com uma rede de superfície. In: **XX Simpósio Brasileiro de Recursos Hídricos**. [S.l.: s.n.], 2013.

GOURLEY, J. J.; CALVERT, C. M. Automated detection of the bright band using WSR-88D data. **Weather and Forecasting**, v. 18, p. 585–599, 2003.

GREENE, D. R.; CLARK, R. A. Vertically integrated liquid water - a new analysis tool. **Monthly Weather Review**, v. 100, n. 7, p. 548–552, 1972.

ILLINOIS. **University of Illinois WW2010 Project**. Fev 2013. Disponível em: <<http://ww2010.atmos.uiuc.edu/%28Gh%29/guides/rs/rad/basics/angl.xml>>.

INMET. **Instituto Nacional de Meteorologia**. Fev 2014. Disponível em: <<http://www.inmet.gov.br>>.

JUNIOR, N. L. C. **Clusterização baseada em algoritmos fuzzy**. Dissertação (Mestrado) — Programa de Pós-graduação em Ciência da Computação, Universidade Federal de Pernambuco, 2006.

MARSHALL, J. S.; PALMER, W. M. The distribution of raindrops with size. **Journal of Meteorology**, v. 5, p. 165–166, 1948.

METEOPT. **Fórum de meteorologia**. Dez 2013. Disponível em: <<http://www.meteopt.com/>>.

MITTERMAIER, M. P.; ILLINGWORTH, A. J. Comparison of model-derived and radar-observed freezing-level heights: Implications for vertical reflectivity profile-correction schemes. **Quarterly Journal of the Royal Meteorological Society**, v. 129, p. 83–95, 2003.

NETO, M. A. S. **Mineração visual de dados: extração do conhecimento a partir das técnicas de visualização da informação e mineração de dados**. Dissertação (Mestrado) — Programa de Pós-Graduação em Métodos Numéricos para Engenharia, Universidade Federal do Paraná, 2008.

PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

PINHEIRO, L. C.; VAZ, M. S. M. G.; MARTINHAGO, A. Z. Proposta de uma extensão do padrão FGDC/CSGDM para dados de radar meteorológico. **Publicatio UEPG, Ciências Exatas e da Terra, Ciências Agrárias e Engenharias**, v. 11, n. 3, p. 33–44, 2005.

QUEIROZ, A. P. de. **Monitoramento e previsão imediata de tempestades severas usando dados de radar**. Dissertação (Mestrado) — Curso de Pós-Graduação em Meteorologia, Instituto Nacional de Pesquisas Espaciais, 2009.

RINEHART, R. E. **Radar for meteorologists**. 4. ed. Nevada, United States of America: Rinehart Publications, 2004. ISBN 0-9658002-1-0.

SANCHEZ-DIEZMA, R.; ZAWADZKI, I.; SEMPERE-TORRES, D. Identification of the bright band through the analysis of volumetric radar data. **Journal of Geophysical Research**, v. 105, n. D2, p. 2225–2236, 2000.

SANTOS, A. L. B. A. dos. **Uma nova metodologia para a recuperação do campo de vento**. Dissertação (Mestrado) — Programa de Pós-Graduação em Métodos Numéricos para Engenharia, Universidade Federal do Paraná, 2011.

SANTOS, T. N. dos. **Redes neurais artificiais e relação ZR aplicadas à estimativa de chuva**. Dissertação (Mestrado) — Programa de Pós-Graduação em Métodos Numéricos para Engenharia, Universidade Federal do Paraná, 2014.

SIMEPAR. **Sistema Meteorológico do Paraná**. Fev 2014. Disponível em: <<http://www.simepar.br>>.

SMITH, C. J. The reduction of errors caused by bright bands in quantitative rainfall measurements made using radar. **Journal of Atmospheric and Oceanic Technology**, v. 3, p. 129–141, 1986.

SMYTH, T. J.; ILLINGWORTH, A. J. Radar estimates of rainfall rates at the ground in bright band and non-bright band events. **Quarterly Journal of the Royal Meteorological Society**, v. 124, p. 2417–2434, 1998.

SOUZA, T. A. de. **Agrupamento de séries temporais de vento para avaliação da disponibilidade de geração de usinas eólicas**. Dissertação (Mestrado em Engenharia Elétrica) — Programa de Pós-Graduação de Engenharia, Universidade Federal do Rio de Janeiro, 2008.

THEODORIDIS, S.; KOUTROUMBAS, K. **Pattern recognition**. 4. ed. [S.l.]: Academic Press, 2009. ISBN 978-1-59749-272-0.

VIGNAL, B. et al. Three methods to determine profiles of reflectivity from volumetric radar data to correct precipitation estimates. **Journal of Applied Meteorology**, v. 39, p. 1715–1726, 2000.

ZHANG, J.; LANGSTON, C.; HOWARD, K. Brightband identification based on vertical profiles of reflectivity from the WSR-88D. **Journal of Atmospheric and Oceanic Technology**, v. 25, p. 1859–1872, 2008.

ZHANG, J.; QI, Y. A real-time algorithm for the correction of brightband effects in radar-derived QPE. **Journal of Hydrometeorology**, v. 11, p. 1157–1171, 2010.